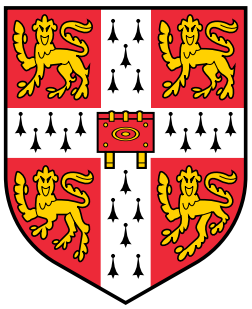
Linear-Scaling Density Functional Theory and Theoretical Electron Energy Loss Spectroscopy Investigations of Surfaces and Defects in Nanomaterials

University of Cambridge



Peterhouse

This dissertation is submitted for the degree of Doctor of Philosophy

Edward William Tait

June 2018

Contents

Abstract

This work presents a method for the computation of electron energy loss spectra for large systems using the ONETEP code.

The foundations of density functional theory will be discussed, with a focus on the linear scaling methods employed by ONETEP. A method for the prediction of spectra will be shown, with emphasis on the construction of the matrix elements needed. This method is tested against both other DFT codes and experiment.

Once the ability of ONETEP to predict EEL spectra has been established the code is used to predict the spectra associated with oxygen vacancy defects in the anatase (101) surface. These spectra are found to agree with the limited experimental data available. EEL spectra for nitrogen dopants in anatase are also predicted, based on literature geometries for these.

The energetics of the oxygen vacancy is also explored, using the ability of ONETEP to perform calculations on large cells to address elastic finite size effects and employing a correction scheme to remove electrostatic finite size effects.

Finally, the properties of an extended grain boundaries defect in GaAs were investigated, focussing on properties relevant for optoelectronics.

Dedication

For Kit & Bill

Declaration

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text.

This dissertation is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my dissertation has already been submitted, or, is being concurrently submitted for any such degree, diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

This dissertation does not exceed 60,000 words.

Edward Tait

Acknowledgements

I would like to thank my supervisors Nick Hine and Caterina Ducati for their ongoing advice. Without the assistance of Michael Rutter my coding would have proven far more difficult and the end result would be much less efficient.

This work was made possible by the generous funding of the EPSRC via the Cambridge Nanotechnology Doctoral Training Centre (NanoDTC).

Finally I would must thank Geoff, who has been with me through thick and thin and after all that agreed to proof read.

Preface

The work in chapter 7 was undertaken in collaboration with the experimental groups of A. Sanchez and R. Beanland at the University of Warwick. These groups undertook synthesis and microscopy of the GaAs nanomaterials while I performed computational studies.

Chapter 1

Introduction

1.1 Motivation

1.1.1 Defects in Materials

Designing novel functional materials requires a detailed knowledge of the relationship between the structure of a material and its properties. Defects are one structural feature which play a profound role in controlling the properties of materials. While in many situations defects may be unwelcome, in others they are essential to desired functions. For example, doping in semiconductors is based upon substitutional defects, while many heterogeneous catalysts derive their activity from surface defects where substrates can attach.

Without understanding which defect species are present in a material and the effects those defects can have, control of material properties is unlikely to be achieved. Developing a tool to aid in determining which defects are present in a material is the objective of this work. Methods of studying the structure of defects using a combination of theory and experiment will be discussed. This is followed by work on the formation energy of a specific defect.

1.2 Transition Metal Oxides

Transition metal oxides posses a range of behaviours and properties as diverse as the metals they derive from. The first row metals are typically abundant, leading to low prices and if not entirely benign tend not to be as severely toxic as heavy metals.

Present applications often exploit both passive and active properties. TMOs find use as catalysts for a number of industrial processes (the contact process for example). Other 'active' applications include in electronics, either as gate insulators in field effect transistors, or in metal oxide varistors - used to protect against power surges. In more passive roles the stability of these materials lends them to use in pigments and fillers.

Looking to the future, roles involving photovoltaic and photocatalytic activity are envisaged. In these particular roles, one TMO has risen to particular prominence: Titanium dioxide, which benefits from all of the advantages listed earlier.

1.2.1 Titanium Dioxide

Titanium dioxide is a commonly encountered material. Its low toxicity has enabled it to displace lead based white pigments in paint. It is sufficiently non toxic to find use in food and pharmaceuticals. Anatase is one of two major polymorphs of TiO_2, the other being Rutile. Anatase is of particular interest owing to its promising performance in the areas outlined above. Anatase has been shown to effect the photodegradation of various organic compounds, including methylene blue [1] amongst other dyes [2], chlorophenols [3] and chemical weapons [4]. The photocatalytic properties of anatase have found commercial applications in self-cleaning window glass and in paving intended to reduce air pollution. Typically, various doping schemes are used in an attempt to increase activity. The objective is to reduce the band gap of anatase to energies around the peak of the solar spectrum. Commonly used

dopants include nitrogen [5], sulphur [6] and carbon [7]. Numerous groups are working on photochemical water splitting using an anatase photocatalyst. The objective here is to produce solar derived fuels (typically H_2).

Interest in direct photovoltaic applications of TiO_2 was aroused initially by the application of nanocrystaline TiO_2 in dye sensitised solar cells, where TiO_2 nano-particles are coupled to an organic or organometallic sensitizer which enables the capture of photons with an energy less than the band gap of TiO_2. The discovery of organic perovskite technology had displaced DSSCs somewhat, but these new devices still sometimes use TiO_2 as a support and charge-collecting material. Both these technologies have the potential to revolutionise solar energy production, as they do not depend on expensive conventional semiconductor fabrication methods to produce.

1.3 Experimental Characterisation

Determining the structure of complex materials is a major experimental challenge. The structure of bulk crystalline solids may be elucidated using techniques such as x-ray diffraction which, when successful, provide complete structural information.

If a material has a complex nanostructure, diffraction methods may prove insufficient. Depending on feature size, scanning electron microscopy, atomic force microscopy, scanning tunnelling microscopy and transmission electron microscopy may all be of use. Given defects tend to be on the scale of a few atoms in at least one dimension, the last of these methods is the most relevant here.

In transmission electron microscopy, a beam of electrons is passed through the sample and electron lenses are used to form a greatly magnified image of the sample on a detector, initially a fluorescent screen but latterly CCDs have become popular. Magnifications of several hundred thousand fold are possible and individual atomic positions may be resolved. Recent improve-

ments in instruments enable ever greater resolution. This is because electron microscope resolution is not limited by the wavelength of the probe (typically less than 0.1Å) but instead by the quality of the electron optics[8]. Advances in computer control have enabled the development of aberration correctors, which counteract spherical aberration. Further more, development in electron gun technology has enabled the probe beam to have a much lower energy spread. This lower energy range reduces chromatic aberration and also benefits spectroscopic methods.

1.3.1 Electron Energy Loss Spectroscopy

One experimental technique frequently used in the analysis of nano-materials is Scanning Transmission Electron Microscopy (STEM). As well as providing morphological information, STEM can give a wealth of additional data in the form of spectroscopic measurements. Electron energy loss spectroscopy (EELS) is one such spectroscopic measurement. Here, the energy loss of the scanning electron beam is measured as it passes through the sample[9]. Combination with imaging means that a site of particular interest can be located prior to spectroscopy and spectral maps of an image can be compiled, with an EEL spectrum for each pixel. The advances in electron guns described above have led to a progressive improvement in energy resolution in EEL spectra, with values as low as 200 meV now being obtained when Wien Filter monochromators are used [10]. The EEL spectrum contains a great deal of information about a material. Some key features include: the background, which decays with increasing energy. This can be removed in post processing steps. At around the zero point on the spectrum are energy losses associated with the excitation of phonons, plasmons, other low energy collective excitations (such as magnons) if the material supports them and transitions from the valence manifold to the conduction band. This is called the low loss region.

At higher energies, abrupt increases in count rate are seen. A schematic

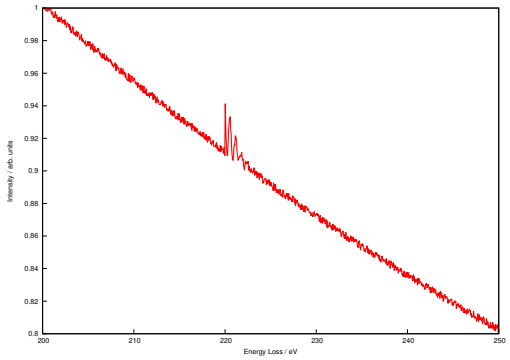


Figure 1.1: A schematic of a typical EEL spectrum.

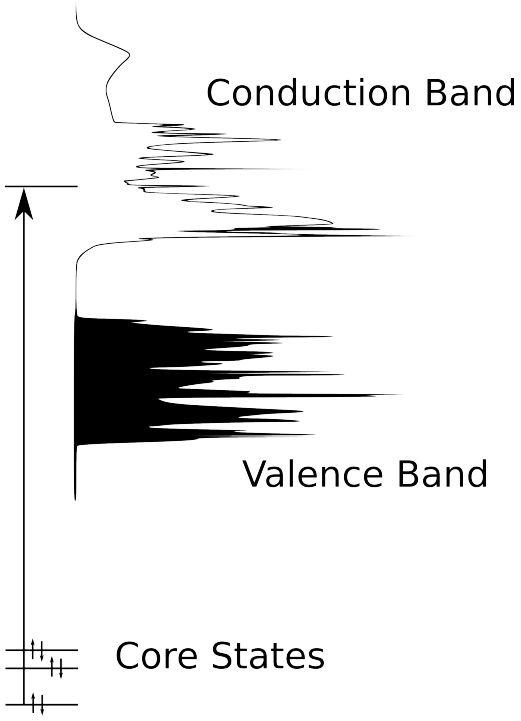


Figure 1.2: The transitions involved in core loss EELS are from 'core' orbitals into the conduction band of the sample.

example of this region of a spectrum is shown in figure 1.1. Note the decaying background and noise, as well as the abrupt onset of the feature. Characteristic 'Edges' correspond to the excitation of electrons from core orbitals to the conduction band of the sample (depicted in Figure 1.2). This region of the spectrum is the *core loss* region and it is the focus of this work. The edges are labelled using the notation from X-ray spectroscopy, with transitions from the 1s orbital being labelled K and so on. Gross energy offset of these edges provides information about which elements are present. This is not all that can be determined. The fine structure of the edge depends on the local coordination and oxidation state of the excited atom, thus local structural information is available. The most common use of EELS within the STEM at present is to produce maps of a sample showing where particular elements are present. This enables the identification of boundaries between materials, impurities and segregation.

Some examples of what has been accomplished using EELS include elemental mapping in a vast range of systems including nanostructures [11–13] and interfaces [14, 15]. Information from near edge structure has been used to identify phases [16–18], Van Hove singularities in the electronic density of states [19] and, combined with *ab initio* methods, determine structures which could not be studied directly using more conventional methods [20].

1.4 The Utility of Theory

While electron energy loss spectroscopy is a powerful experimental technique, it can be difficult to go from an EEL spectrum back to the structure of the material. Here theory can be of assistance. EEL spectra for several possible structures can be simulated and the results compared to the experimental spectrum.

The objective of this work is to develop an implementation of EELS within the ONETEP code. By combining EELS simulation with linear scaling DFT,

it is hoped that larger and more complex nano structures may be analysed and previously non-interpretable spectra may be explained.

1.5 Existing Methods of EELS Simulation

Two main methods currently exist for the simulation of EEL spectra. These approach the problem from two distinct avenues.

1.5.1 Multiple Scattering Codes

These methods rely on a Green's function approach. In this formalism the behaviour of the promoted electron is modeled in terms of a series of scattering processes of an excited electron. Ultimately the loss function is expressed in terms of matrix elements of the form

$$-\frac{1}{\pi}\Im\langle i|\epsilon^* \cdot \mathbf{r}\mathrm{G}(\mathbf{r},\mathbf{r}';E)\epsilon \cdot \mathbf{r}'|i\rangle\,, \tag{1.1}$$

where $\Im$ indicates the imaginary part of the subsequent expression. $|i\rangle$ is an initial core state, $\epsilon \cdot \mathbf{r}$ is the dipole operator representing the perturbing influence of the electron beam and G is the Green's function or *propagator*.

The Green's function may then be expressed in terms of an unperturbed electron propagator and a term describing the electron's scattering interaction with nearby atoms. A series expansion then allows the Green's function to be expressed in terms of single, double, triple etc scattering events. The most popular code of this type is FEFF [21, 22].

Whilst this method is appealing, generally the scattering potentials used are fixed. Thus, any complex behaviour arising out of the specific sample geometry is missed. In the study of defects where local electronic structure may vary significantly, this is a severe drawback.

1.5.2 Electronic Structure Codes

These consider the problem of computing an EEL spectrum as one of computing the energies and relative intensities of transitions between core and conduction band states. In these implementations the ground state electron structure and the low lying conduction band states are first calculated, then the EEL spectrum is simulated as a post processing step.

Many implementations exist, both in all electron and pseudopotential DFT codes. These methods are not limited to DFT, as any technique which produces wavefunctions for core and conduction band states could be used.

As electronic structure codes provide material-specific local information about a material, this is the approach which is used going forward in this work.

A Brief History of EELS Prediction In Electronic Structure Codes

The fundamental basis for EELS simulation within this formalism was laid down by Bethe [23]. Leapman used this result to derive analytic expressions for differential scattering cross sections in boron nitride, which matched well with experimental results for the angular dependence of core edges observed in that material [24].

As electronic structure codes grew more advanced with increasing computer power, the ability to predict core edge structure was added to a number of codes. Initially this functionality was implemented in all electron DFT codes, as these provided access to both core and conduction wavefunctions which were unencumbered by the effects of pseudopotentials. Augmented plane-wave and spherical-wave codes, similar to many of today's all electron codes, were used. Even shortly after the development of these methods it was apparent that they were ill suited to tackling calculations on larger supercells [25].

Ultimately, pseudopotential based methods were developed which enabled

the calculation of edge structure in far larger systems. These methods are discussed in more detail in subsequent chapters.

The current state of the art in EELS simulation, at least from an electronic structure perspective, has moved on to include not just single particle methods like DFT but also many particle techniques including Bethe-Salpeter equation and configuration interaction. The former has been used to address core hole interactions directly, without resort to the somewhat *ad hoc* methods used in single particle codes [26]. The later has provided a means of simulating transition metal $L_{2,3}$ edges, in which electron localisation must be considered [27]. A review of this area has been produced by Mizoguchi *et al.* [28].

1.6 Structure

Having demonstrated the importance of defects in semiconductor materials this work will now move on to a discussion of the theoretical methods used. This discussion of density functional theory will establish the fundamentals of the technique before introducing more advanced topics which are required in subsequent chapters. Once DFT has been introduced, the implementation of EELS in the ONETEP code will be discussed. This chapter deals both with the theoretical underpinnings of EELS simulation and the details of the implementation in ONETEP. Comparisons to both existing implementations and experiment are made, taking in a variety of different systems. Having shown that the EELS simulation functionality works as expected, the next two chapters seek to address one of the key aims of this thesis - the application of theoretical EELS to defects in semiconductors. Important defects in anatase are selected for study: the oxygen vacancy and the nitrogen dopant. Key modifications to the spectrum of anatase arising from these defects are identified. Attempts to assess the concentrations at which defects might be detectable to experiment are made.

Being able to identify a defect spectroscopically is not the only way in which theory can help. In Chapter 6 the problem of defect formation energy is discussed. The anatase surface presents a great challenge for the study of charged defects: the dielectric is both anisotropic and inhomogeneous. To deal with this complexity, a series of schemes are introduced and analysed. The most suitable scheme is then implemented and used to correct finite size interactions for charged defects in the (101) anatase surface. The prime objective of the work is to show how defect formation energy converges towards the bulk value as the depth of a defect into the surface is increased. In the course of the investigation, it became clear that oxygen vacancy states in anatase, except in their highest charge state, are not well localised. The problems arising from this are also discussed at length.

While transition metal oxides are the main focus of this work, it is important to understand that the methods discussed here are generally applicable. To this end, a study is made of a different semiconducting system: Gallium Arsenide nanowires. This investigation was motivated by experimental observations. A novel extended defect was observed in the TEM. While the structure could readily be determined from the micrographs, the impact on band structure was not. Simulations showed that this defect is expected to have a substantial impact on the optoelectronic properties of GaAs nanowires.

Chapter 2

Density Functional Theory

2.1 Introduction

All methods of electronic structure calculation offer a trade-off between speed and accuracy. With presently available technology, density functional theory (DFT) is widely considered to offer a good balance between the two.

The first part of this chapter will provide an introduction to the physical underpinnings of DFT, proofs of its core theorems and some discussions of their implications. The second half of the chapter is dedicated to some more specific topics needed for understanding the methods developed in later chapters. The projector augmented wave approach for handling atomic potentials is introduced and the topic of linear scaling DFT is covered. Particular attention is given to the code used in this work, with a review of its theoretical underpinnings and unique features.

2.1.1 Some Notes On Notation

Throughout this chapter, atomic units ($e = \hbar = m_e = 1$) have been used. From time to time, the asymptotic scaling behaviour of tasks is discussed. The standard notation is used here. To say a task scales as $\mathcal{O}(f(n))$ indicates

that there exists a finite c such that as n becomes large the amount of a resource (computational time $\propto$ operations unless stated otherwise) required to perform that task is bounded from above by $cf(n)$.

2.1.2 Hohenberg-Kohn

The great advantage of DFT is that it removes the need to directly solve the interacting many body Schrödinger equation, working instead with the electron density of the system. The method relies on two results due to Hohenberg and Kohn [29]:

1. The ground state electron density of a system uniquely defines its Hamiltonian (up to an additive constant).

2. There exists a *Universal Functional*, $E[n]$, which is minimised for a specific external potential by the exact ground state density for a system with that potential.

An important result follows form the first of these theorems: since the ground state density defines the Hamiltonian of a system, it must also define the eigenstates of the system. Consequently, *any property* of a system can *in principle* be determined from its ground state density.

Proof of the First Hohenberg-Kohn Theorem

Let us assume two Hamiltonians, H_1 and H_2, produce the same ground state density n and differ by some potential ΔV. The two Hamiltonians must obey their respective Schrödinger equations. The solutions to these equations are $\Psi(\mathbf{r}_1 \dots \mathbf{r}_\mathrm{N})$: all-electron many body wave functions depending on the electron coordinates $\mathbf{r}_1 \dots \mathbf{r}_\mathrm{N}$ of the N electrons in the system. Here spin has been ignored for clarity.

$$H_1\Psi_1(\mathbf{r}_1 \dots \mathbf{r}_\mathrm{N}) = E_1\Psi_1(\mathbf{r}_1 \dots \mathbf{r}_\mathrm{N}) \tag{2.1}$$

$$H_2\Psi_2(\mathbf{r}_1 \dots \mathbf{r}_\mathrm{N}) = E_2\Psi_2(\mathbf{r}_1 \dots \mathbf{r}_\mathrm{N}) \tag{2.2}$$

$$\begin{aligned} &\int |\Psi_1(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_\mathrm{N})|^2 \mathrm{d}\mathbf{r}_2 \ldots \mathrm{d}\mathbf{r}_\mathrm{N} \\ =&\int |\Psi_2(\mathbf{r}, \mathbf{r}_2, \ldots, \mathbf{r}_\mathrm{N})|^2 \mathrm{d}\mathbf{r}_2 \ldots \mathrm{d}\mathbf{r}_\mathrm{N} = n(\mathbf{r}) \end{aligned} \tag{2.3}$$

The proof will use the variational theorem. The expected energy of a trial wavefunction will be greater than or equal to that of the true ground state wavefunction of a system, with equality only being achieved in the case where the trial wavefunction is equal to the exact ground state wavefunction. Thus for a trial wave function $\widetilde{\Psi} \neq \Psi_0$ and Hamiltonian H with ground state energy E_0 given by $H\Psi_0 = E_0\Psi_0$:

$$\langle \widetilde{\Psi} | H | \widetilde{\Psi} \rangle > E_0 \,. \tag{2.4}$$

Substituting H_1 and Ψ_2 and *vice versa* into Eq. 2.4 yields:

$$\begin{aligned} \langle \Psi_2 | H_1 | \Psi_2 \rangle > E_1 \,, \\ \langle \Psi_1 | H_2 | \Psi_1 \rangle > E_2 \,. \end{aligned} \tag{2.5}$$

Since the two Hamiltonians describe the same number of electrons, the only way they can differ is in the external potential. We write this difference as ΔV:

$$H_2 = H_1 + \Delta V \,. \tag{2.6}$$

Using Eq. 2.6 to substitute into the two inequalities Eq. 2.5 gives:

$$\begin{aligned} \langle \Psi_2 | H_2 - \Delta V | \Psi_2 \rangle > E_1 \,, \\ \langle \Psi_1 | H_1 + \Delta V | \Psi_1 \rangle > E_2 \,. \end{aligned} \tag{2.7}$$

Noting that $\langle \Psi_1 | H_1 | \Psi_1 \rangle = E_1$ and likewise for Ψ_2 and H_2, it is possible to simplify to:

$$\begin{aligned} E_2 - \langle \Psi_2 | \Delta V | \Psi_2 \rangle > E_1 \,, \\ E_1 + \langle \Psi_1 | \Delta V | \Psi_1 \rangle > E_2 \,. \end{aligned} \tag{2.8}$$

Using Eq. 2.3 and the definition of the braket we find:

$$\langle \Psi_2 | \Delta V | \Psi_2 \rangle = \int \mathrm{d}\mathbf{r} n(\mathbf{r}) \Delta V(\mathbf{r}) = \langle \Psi_1 | \Delta V | \Psi_1 \rangle = \langle \Delta V \rangle \,. \tag{2.9}$$

Simplifying Eq. 2.8 using Eq. 2.9 gives:

$$\begin{aligned} E_2 - \langle \Delta V \rangle &> E_1 \,, \\ E_1 + \langle \Delta V \rangle &> E_2 \,. \end{aligned} \tag{2.10}$$

Adding, we finally obtain the contradictory inequality:

$$E_1 + E_2 > E_1 + E_2 \,. \tag{2.11}$$

The contradiction Eq. 2.11 demonstrates that the initial assumption, that two Hamiltonians may have the same ground state density, was incorrect and thus proves the first Hohenberg-Kohn theorem.

Proof of the Second Hohenberg-Kohn Theorem

The first Hohenberg-Kohn Theorem established that, by defining the Hamiltonian, a ground state density encodes all possible information about a system. It follows that the kinetic and potential energies may be written in terms of the density:

$$\begin{aligned} E_{\mathrm{k}} &= T[n] \\ E_{\mathrm{pot}} &= E_{\mathrm{pot}}[n] \end{aligned} \tag{2.12}$$

It is possible to further split the potential energy term into a contribution from the interactions of the electrons with the external potential and one from the electrons interacting with each other:

$$E_{\mathrm{pot}}[n] = E_{\mathrm{e-e}}[n] + \int n(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r}) \mathrm{d}\mathbf{r} \tag{2.13}$$

Combining the kinetic and potential terms gives a functional for the overall energy:

$$\begin{aligned} E[n] &= E_{\mathrm{k}}[n] + E_{\mathrm{pot}}[n] = T[n] + E_{\mathrm{e-e}}[n] + \int n(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r}) \mathrm{d}\mathbf{r} \\ &= E_{\mathrm{HK}}[n] + \int n(\mathbf{r}) V_{\mathrm{ext}}(\mathbf{r}) \mathrm{d}\mathbf{r} \end{aligned} \tag{2.14}$$

Equation Eq. 2.14 is taken to define $E_{\mathrm{HK}}[n]$. This functional has no explicit dependence on V_{ext} and is applicable to all-electron systems. It is *Universal*.

2.1.3 Kohn-Sham

The results of the previous section are exact. Were the form of $E_{\mathrm{HK}}[n]$ known, it would be possible to compute the ground state energy of a system by finding the $n(\mathbf{r})$ which minimises Eq. 2.14. Unfortunately an expression for neither sub component of $E_{\mathrm{HK}}[n]$ is known in the case of interacting electrons.

To make progress it will be necessary to make approximations, with an *ansatz* from Kohn and Sham has enabled the field to progress. The idea is this: *assume* that the ground state density of the system of interest is also the ground state density of an auxiliary system of non-interacting electrons.

Because the particles in the auxiliary system do not interact, the many body wavefunction of the system may be expressed in terms of single-particle states. Using a Slater determinant ensures that the wavefunction is antisymmetric under particle exchange. Let ψ_i^{KS} be a single particle wavefunction with energy E_i. In the ground state, a system with N electrons will have the lowest N levels filled. We can thus write the many body wavefunction of the auxiliary system as the Slater determinant:

$$\Psi^{\mathrm{KS}}(\mathbf{r}_1, \cdots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1^{\mathrm{KS}}(\mathbf{r}_1) & \psi_2^{\mathrm{KS}}(\mathbf{r}_1) & \cdots & \psi_N^{\mathrm{KS}}(\mathbf{r}_1) \\ \psi_1^{\mathrm{KS}}(\mathbf{r}_2) & \psi_2^{\mathrm{KS}}(\mathbf{r}_2) & \cdots & \psi_N^{\mathrm{KS}}(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1^{\mathrm{KS}}(\mathbf{r}_N) & \psi_2^{\mathrm{KS}}(\mathbf{r}_N) & \cdots & \psi_N^{\mathrm{KS}}(\mathbf{r}_N) \end{vmatrix} \tag{2.15}$$

As stated above, the density of this system can be written:

$$\sum_i |\psi_i^{\mathrm{KS}}(\mathbf{r})|^2 = n(\mathbf{r})\,. \tag{2.16}$$

The energy of this system may be written as:

$$E_{\mathrm{KS}}[n] = T_{\mathrm{KS}}[n] + E_{\mathrm{H}}[n] + \int n(\mathbf{r}) v_{\mathrm{ext}}(\mathbf{r}) \mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[n]\,, \tag{2.17}$$

with:

$$T_{\mathrm{KS}}[n] = -\frac{1}{2}\sum_{i}^{N}\int \psi_i^{\mathrm{KS}*}(\mathbf{r})\nabla^2\psi_i^{\mathrm{KS}}(\mathbf{r})\mathrm{d}\mathbf{r}\,, \tag{2.18}$$

$$E_{\mathrm{H}}[n] = \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\,, \tag{2.19}$$

$$v_{\mathrm{ext}}(\mathbf{r}) = -\sum_{j\in\mathrm{Nuclei}}\frac{Z_j}{|\mathbf{r}-\mathbf{R}_j|}\ \text{(Coulomb interaction with nuclei)}\,. \tag{2.20}$$

The final term in Eq. 2.17 is the exchange correlation functional. We can form a connection to the universal functional of the second Hohenberg-Kohn theorem and $E_{\mathrm{xc}}[n]$. The total energy associated with the density n has been expressed in two ways which must be equal:

$$\begin{aligned} T_{\mathrm{KS}}[n] + E_{\mathrm{H}}[n] + \int n(\mathbf{r})v_{\mathrm{ext}}(\mathbf{r})\mathrm{d}\mathbf{r} + E_{\mathrm{xc}}[n] = \\ T[n] + E_{\mathrm{e-e}}[n] + \int n(\mathbf{r})V_{\mathrm{ext}}(\mathbf{r})\mathrm{d}\mathbf{r} \end{aligned} \tag{2.21}$$

We may simplify this expression by cancelling the external potential term and rearrange to express E_{xc} in terms of other quantities:

$$\begin{aligned} E_{\mathrm{xc}}[n] &= T[n] + E_{\mathrm{e-e}}[n] - (T_{\mathrm{KS}}[n] + E_{\mathrm{H}}[n]) \\ &= (T[n] - T_{\mathrm{KS}}[n]) + (E_{\mathrm{e-e}}[n] - E_{\mathrm{H}}[n]) \end{aligned} \tag{2.22}$$

From Eq. 2.22 we can understand why the contribution from the exchange-correlation functional tends to be small. $E_{\mathrm{xc}}[n]$ contains two contributions. The first accounts for the difference in kinetic energy between the interacting and non-interacting system and the second for the difference in potential energy. In as far as the many body KS wavefunction represents the wavefunction of the interacting system, these differences will be small and thus $E_{\mathrm{xc}}[n]$ is small.

The Kohn-Sham Equations

At this point the task of finding the total energy has been changed from finding the density which minimises $E[n]$ to finding the $\{\psi_i^{\mathrm{KS}}\}$ which minimise $E_{\mathrm{KS}}[n]$.

This task is not as daunting as it first appears. Applying the calculus of variations with a Lagrange multiplier to ensure constant particle number yields, after some algebra, the Kohn-Sham equations:

$$\left(-\frac{1}{2}\nabla^2 + v_{\text{ext}}(\mathbf{r}) + v_{\text{h}}(\mathbf{r}) + v_{\text{xc}}(\mathbf{r})\right)\psi_i^{\text{KS}} = \epsilon_i \psi_i^{\text{KS}} \tag{2.23}$$

The external potential in Eq. 2.23 has its usual definition. The remaining two are defined as

$$v_{\text{h}}(\mathbf{r}) = \frac{1}{2}\int \mathrm{d}\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\,, \tag{2.24}$$

$$v_{\text{xc}}(\mathbf{r}) = \frac{\delta E_{\text{xc}}[n]}{\delta n(\mathbf{r})}\,. \tag{2.25}$$

Equations 2.23 resemble the Schrödinger equation for a non-interacting particle, which feels both the external potential from the nuclei, v_{ext} but also the Hartree potential, arising from a classical Coulomb interaction with the electron density, v_{h} and the exchange correlation potential, v_{xc}, which accounts for the quantum interactions of the electrons in the physical system. The resulting system of non-linear equations must be solved iteratively until a self-consistent (up to some tolerance) state is reached. In order to solve these equations, v_{xc} must be known. Unfortunately an exact form for the exchange-correlation potential remains elusive and approximations must be used.

2.1.4 Exchange Correlation Functionals

There are a wide variety of functionals described in the literature. The LibXC project, for example implements over 400 different exchange and correlation functionals [30]. Often these are designed and tested with the intent of providing accurate results in some class of situations; for example in solids or reproduction of Van der Waals interactions. Others attempt to be generalists, doing a reasonable job for a diverse range of systems.

Functionals are sometimes thought of as being arranged in a *Jacob's Ladder*, with simpler and supposedly less accurate functionals on the lower rungs and more complex functionals rising towards the heaven of the exact exchange correlation functional[31].

On the lowest rung are the local density approximations (originally proposed by Kohn & Sham [32]), which perform well considering their simplicity. To construct an LDA functional, the functional is split into exchange and correlation parts:

$$E_{\mathrm{xc,LDA}}[n] = E_{\mathrm{X}}[n] + E_{\mathrm{C}}[n]\,. \tag{2.26}$$

For a homogeneous electron gas, the exchange energy is known [33] exactly to be:

$$E_{\mathrm{x}}[n] = -\frac{3}{3}(\frac{\pi}{3})^{\frac{1}{3}} \int n(\mathbf{r})^{\frac{4}{3}} \mathrm{d}\mathbf{r}\,. \tag{2.27}$$

The correlation functional requires more work. Fits [34] to quantum Monte-Carlo simulations [35] have been undertaken, with parametrisations typically providing 0.1mHa agreement with QMC over several orders of magnitude of electron density. The diversity in LDAs comes largely from the choice of parametrisation for $E_c[n]$.

One explanation for the unexpected accuracy of the LDA is that is satisfies the sum rule: that exactly one electron should be excluded by the exchange correlation hole. The exact functional is known to satisfy this requirement but several more sophisticated approximations do not [36].

On the next rung sit the generalised gradient approximations (GGAs). These seek to tackle the wholly local nature of the LDAs by including terms such as the gradient of the density. In this way, the effects of inhomogeneity may be directly addressed. The inclusion of additional dependence greatly increases the number of possible formulations available and GGA functionals abound. Some common examples include the functional of Perdew, Burke and Ernzerhof - PBE [37] and its many variants and improvements [38–40]. Some GGAs do not obey conditions the exact functional is known to obey.

Other quantities can the added to the set used in producing the functional, including the kinetic energy density results in the meta-GGAs such as the functional of Tao, Perdew, Staroverov, and Scuseria [41].

Noting that the KS wavefunctions are implicit functionals of the density, there are those exchange correlation functionals which use them. Notably use of the KS eigenstates permits computation of exact exchange which can either be used on its own (reducing DFT to Hartree-Fock) or combined with some weight with another functional to produce a hybrid functional - the possibilities are effectively limitless but come with the caveat of being both rather empirical and significantly more computationally costly than (semi)local functionals. B3LYP [42] is a classic example of this class of functional.

Finally, in addition to including occupied orbitals, it is also possible to include unoccupied orbitals, this leads to random phase approximation type methods.

For most of this study, the plain PBE functional [37] was selected as it is well known and tested. In Chapter 6, some studies were conducted using the HSE06 [43–45] hybrid functional to investigate the impact of self-interaction errors on the tendency of electron density to migrate to the surface of a system.

2.1.5 Projector Augmented Wave

Close to the nucleus of an atom, the all-electron wavefunctions oscillate rapidly (inter-nodal distances on the order of a Bohr or less). Representing such wavefunctions with plane waves is undesirable, as a very large basis would be required making all but the smallest calculations computationally infeasible.

To address this issue, pseudopotentials were introduced. Chemical intuition indicates that core electrons play no part in bonding and thus need not be explicitly included in a simulation which is interested in computing

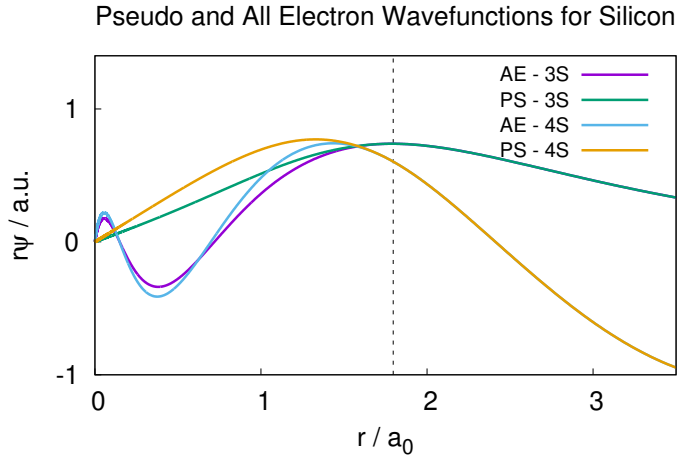


Figure 2.1: Pseudo and all-electron S wavefunctions, expressed as $r\psi(r)$ for silicon. The vertical dashed line indicates the core radius, outside which the pseudo and all-electron wavefunctions must agree. Note that the pseudo wavefunctions lack the so called 'orthogonality wiggles' inside the core region which would normally require many basis functions to represent. AE $\rightarrow$ All Electron, PS $\rightarrow$ Pseudo.

chemical properties. Furthermore chemical interactions take place in the space between atoms, thus the behaviour of wavefunctions at the nucleus is less important for bonding than might be supposed. The reason for this was explored Cohen and Heine, and shown to arise from a cancellation of potential and kinetic energies for valence wavefunctions in the region of the nucleus [46].

A pseudopotential will seek to correctly reproduce the behaviour of the wavefunction outside some cutoff radius, but within that radius (near the nucleus) the wavefunction is much smoother. Figure 2.1 shows a pseudo vs all-electron radial wavefunction for silicon, as generated by CASTEP.

Modern pseudopotentials may be divided into two broad classes: the norm conserving and the ultrasoft. Norm conserving pseudopotentials (NCPPs)

ensure that the probability of finding an electron within the cutoff radius of the nucleus is the same as in the all-electron atom. The development of NCPPs was a great step forward, as they had improved *transferability* compared previous potentials. That is, they produced accurate results even when used to model systems different from those they were initially calculated for [47, 48]. Denteneer and Van Haeringen provide a classic review of the technique [49].

The norm conservation constraint tends to lead to 'harder' pseudopotentials which can still vary quite rapidly in space and thus require larger basis sets than might be desirable.

Due to the problems with NCPPs, Vanderbilt introduced 'ultrasoft' PPs [50]. These relax the norm conservation condition and thus produce softer potentials. The reduced basis set size which comes from this is not free however, and is paid for with increased implementation complexity. Despite this the speedups and transferability offered by ultrasofts mean they now dominate the 'market'.

For the simulation of core-loss spectra, pseudopotentials present a problem because conduction band wavefunctions must be accurately represented near the nucleus if transition matrix elements are to be correctly calculated. However, this region is exactly the area where wavefunctions obtained using pseudopotentials differ most from all-electron wavefunctions. A solution to this problem is due to Blöchl [51]. In effect, the inaccurate pseudised part of a wave function is subtracted off and the correct all-electron function added back in.

To see how this works we introduce an operator to convert pseudo to all-electron wavefunctions — the $\mathcal{T}$ operators are only non-zero inside their respective core regions, R (Equation 2 from [51]).

$$\mathcal{T} = 1 + \sum_R \mathcal{T}_R \tag{2.28}$$

Inside the core region the effect of $\mathcal{T}_R$ is to transform a basis of 'pseudo partial

waves' into one of 'all-electron partial waves' which are orthogonal to the core states but form an otherwise complete basis:

$$\mathcal{T}|\widetilde{\varphi}_i\rangle = (1 + \mathcal{T}_R)|\widetilde{\varphi}_i\rangle = |\varphi_i\rangle \tag{2.29}$$

The φ_i are an orthogonal set of functions obtained by radial integration of the Schrodinger equation for a set of energies ϵ_i. Specifically the φ_i solve the equation:

$$(-\frac{1}{2}\nabla^2 + v_{\text{at}}(r))|\varphi_i\rangle = \epsilon_i^1|\varphi_i\rangle \tag{2.30}$$

with $v_{\text{at}}(r)$ an atomic potential and ϵ_i^1 an energy which can be chosen to optimise the properties of the partial wave. In practice ϵ_i^1 is often the lowest bound energy for a given angular momentum channel.

The $\widetilde{\varphi}_i$ are defined in an analogous way, using the pseudised Schrodinger Equation instead of the all electron one, v_{at} is replaced with a pseudised form w_i which is constructed to exactly match the all electron potential outside some cut off radius r_k:

$$(-\frac{1}{2}\nabla^2 + w_i(r))|\widetilde{\varphi_i}\rangle = \epsilon_i^1|\widetilde{\varphi_i}\rangle \tag{2.31}$$

with

$$w_i(r) = \tilde{v}_{\text{at}}(r) + c_i k(r) \tag{2.32}$$

where $\tilde{v}_{\text{at}}(r)$ is a suitably chosen atomic pseudopotential and

$$k(r) = \exp\{-\Big(\frac{r}{r_k}\Big)^{\lambda}\} \tag{2.33}$$

with λ determining the sharpness of the transition from the pseudo atomic potential present for $r < r_k$ and the all electron potential present where $r > r_k$. A value of 6 is recommended by [51]. The value of c_i is determined by requiring the pseudo and all electron partial waves are identical for $r > r_k$.

Inside a core region we can expand the pseudo-wavefunction in terms of the pseudo-partial waves:

$$|\widetilde{\psi}^{\text{KS}}\rangle = \sum_i c_i|\widetilde{\varphi}_i\rangle \tag{2.34}$$

Using the definition of $\mathcal{T}$ and Eq. 2.29, we obtain (equation 4 from [51]):

$$|\psi^{\mathrm{KS}}\rangle = \sum_i c_i |\varphi_i\rangle \tag{2.35}$$

Enabling the expression of the all-electron wave function in terms of the pseudo wave function:

$$|\psi^{\mathrm{KS}}\rangle = |\widetilde{\psi}^{\mathrm{KS}}\rangle + \sum_i c_i (|\varphi_i\rangle - |\widetilde{\varphi}_i\rangle) \tag{2.36}$$

Finally, as $\mathcal{T}$ is linear, the coefficients c_i must be derived from $\widetilde{\psi}^{\mathrm{KS}}$ as scalar products with 'projector functions':

$$\begin{aligned} c_i &= \langle p_i | \widetilde{\psi}^{\mathrm{KS}} \rangle \,, \\ \langle p_i | \tilde{\varphi}_j \rangle &= \delta_{ij} \end{aligned} \tag{2.37}$$

The projectors, $\tilde{p}_i$ are computed using an iterative procedure. An initial projector is obtained by applying a modified Hamiltonian to the pseudo partial wave thus:

$$|\tilde{p}_i\rangle = (-\frac{1}{2}\nabla^2 + \tilde{v}_{\mathrm{at}} - \epsilon_i^1)|\tilde{\varphi}_i\rangle \,. \tag{2.38}$$

Once an initial set of projectors has been obtained they are modified to fulfil the orthogonality condition in 2.37.

Combining 2.36 and 2.37 gives (equation 9 from [51]):

$$|\psi^{\mathrm{KS}}\rangle = |\widetilde{\psi}^{\mathrm{KS}}\rangle + \sum_i \langle \widetilde{p}_i | \widetilde{\psi}^{\mathrm{KS}} \rangle (|\varphi_i\rangle - |\widetilde{\varphi}_i\rangle) \tag{2.39}$$

From this expression, the full all-electron expectation value of an operator may be computed using the pseudo wave function, partial waves and projectors, the final two of which may be computed ahead of time.

2.1.6 Basis Functions

In order to solve the Kohn-Sham equations on a computer, the wavefunctions must be represented in some ways. Many options present themselves: the

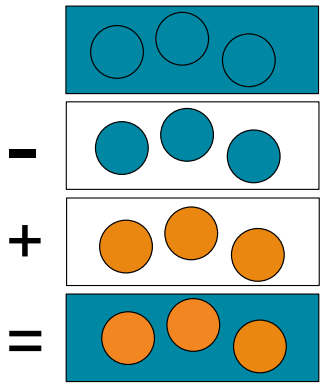


Figure 2.2: Illustration of PAW: At the top of the image the pseudo-wavefunction is shown. Within the augmentation spheres, the pseudo wavefunction (second block) is subtracted off and the all-electron wavefunction (third block) is added back on. This gives the all-electron wavefunction (lowest block).

simplest conceptually is to cover the simulation region with a grid and store the value of ψ_i^{KS} at each grid point. This method is adopted by Octopus [52].

Most codes chose to represent the KS wavefunctions in terms of some other basis functions, typically selected for reasons of computational convenience or physical relevance. Noting that materials are made up of atoms many authors adopt basis sets inspired in some way by atomic orbitals. These *atom centred bases* generally have the advantage of small size. Choices for the actual basis functions vary: Slater type orbitals (ADF) and Gaussians (GAMESS, Gaussian and many others) are typical choices. While these bases have advantages they also suffer from the problem of unequal representation of space: areas far from atoms get worse coverage by basis functions which may lead to lack of transferability, a basis which works well in one system may be inadequate in another. This class of basis set often lacks a systematic way to increase the basis size, which makes converging a calculation with respect to the basis challenging.

Other codes, particularly those dealing with periodic systems use plane

wave bases, which will now be discussed in more detail.

Periodic Systems and Plane Wave Basis Sets

When studying bulk crystalline solids, it is essential to exploit periodicity. Using plane waves as a basis to express the Kohn-Sham orbitals is an easy way to achieve this, with the additional advantages of mathematical simplicity and the availability of efficient algorithms (Fast Fourier Transform) to manipulate functions expressed in this basis and equal treatment of all space, which leads to good transferability across systems.

Bloch's theorem enables a Kohn-Sham wave function to be expressed in terms of a phase factor and a part which is periodic with the lattice:

$$\psi^{\mathrm{KS}}_{\mathbf{k},i}(\mathbf{r}) = e^{\mathrm{i}\mathbf{r}\cdot\mathbf{k}} u_{\mathbf{k},i}(\mathbf{r}), \tag{2.40}$$

Where $\mathbf{k}$ lies in the first Brillouin zone and i is the band index. Expressing the periodic function $u_{\mathbf{k}}$ as a sum over Fourier components, where a finite basis set is used, expressed in terms of a cut off energy:

$$u_{\mathbf{k},i}(\mathbf{r}) = \sum_{\mathbf{G}} C_{\mathbf{G},\mathbf{k},i}\ e^{\mathrm{i}\mathbf{G}\cdot\mathbf{r}}, \quad \frac{1}{2}|\mathbf{G}+\mathbf{k}|^2 < E_{\mathrm{cut}}, \tag{2.41}$$

Where the $\mathbf{G}$ are reciprocal lattice vectors.

As plane waves are orthogonal, increasing the energy cut off (including more $\mathbf{G}$ vectors) will enlarge the basis set used to represent the Kohn-Sham wavefunctions. This increase in available degrees of freedom leads to an improved reproduction of the Kohn-Sham wavefunctions. Applying the variational theorem we see that this improved representation will lead to a monotonic lowering of computed ground state energy towards the minimal value. This has the great advantage of permitting good control of finite basis set error. It is important to remember when performing convergence tests that a quantity of interest may converge slower (or faster) than the total energy with basis size. It is of crucial importance to test that the final result of a simulation, not just the total energy, is well converged.

2.2 Linear Scaling

Conventional DFT codes directly solve the KS equations using a variety of methods. Whilst such codes have lead to a great many advances, they suffer from an inability to tackle systems containing a large number of atoms. This barrier arises because KS wavefunctions must be orthogonal and in order to orthogonalise a wavefunction to a set of other wavefunctions (of size N) one must compute N scalar products. This must be completed for each wavefunction in the set for a total cost of $\mathcal{O}(N^2)$. The final factor of N comes from the cost of computing a scalar product, which is *not* independent of the number of electrons in the system for physically realistic materials. Adding more electrons implies adding more atoms and increasing the size of the simulation cell. When using a local basis set, the first factor will increase the basis set size and when using plane waves the second will. As a result the overall effort of solving the KS equations scales as $\mathcal{O}(N^3)$.

To overcome this scaling barrier several methods have been developed which take advantage of the properties of certain materials or systems to obtain better scaling behaviour [53]. All these approaches must avoid the issue outlined above and thus forego use of KS eigenstates. Approaches are founded on the physical principle that (in insulators at least) matter is 'near sighted' [54, 55], that is changing one part of a system has a vanishingly small effect on some remote part of the system as the separation between those two parts is increased.

A concrete example of this idea is that in insulators the density matrix,

$$\rho(\mathbf{r}, \mathbf{r'}) = \sum_{i}^{N} \psi_i^{\mathrm{KS}*}(\mathbf{r'}) f_i \psi^{\mathrm{KS}}(\mathbf{r}) \,, \tag{2.42}$$

(f_i is the orbital occupation 1 or 0 in an insulator) decays exponentially with $|\mathbf{r} - \mathbf{r'}|$ [56]. This property is used by the ONETEP, BIGDFT [57] and CONQUEST [58] codes, among others. Formulating DFT in terms of the density matrix is useful because the density matrix encodes all information

about the system which is needed to recover expectation values of single particle operators. To obtain the expectation value of an operator using the density matrix, one needs only to take the trace of the product of that operator with the density matrix. For a concrete example consider the total energy of the system which may be written as

$$E_{\text{tot}} = \text{Tr}[\rho \text{H}] - E_{\text{DC}}\,, \tag{2.43}$$

where E_{DC} is a double counting term that accounts for two things: First the trace counts the Hartree energy twice - effectively once from electron a interacting with electron b and then again in the other direction. The second contribution to the double counting term is the handling of exchange-correlation energies. Overall the term has a form like

$$E_{\text{DC}}[n(\mathbf{r})] = \int \text{d}\mathbf{r} n(\mathbf{r})(V_{\text{xc}}(\mathbf{r}) - \epsilon_{\text{xc}}(\mathbf{r}) + \frac{1}{2}v_{\text{h}}(\mathbf{r}))\,. \tag{2.44}$$

Note the first two terms will act to replace the integral of the product of the exchange correlation energy and density with the integral of the exchange correlation energy density and the density. The final term will subtract off the spurious double counting of the Hartree energy described above.

Others use expansion of the Fermi operator [59], a generalisation of the density matrix to finite temperature. This yields a $\mathcal{O}(N^2)$ method but by applying truncation can be run in $\mathcal{O}(N)$ for insulating materials.

Linear scaling in 1D systems and sub-cubic in others may be achieved through pole expansion and selected inversion (PEXSI) [60]. This method has been implemented in the SIESTA and CP2K codes [60, 61].

Divide and conquer and band structure based schemes have also been devised [62]. These methods split the system of interest into a series of subsystems whose density is computed independently. Care is taken to ensure a constant Fermi level across the sub systems. Such a scheme has been implemented in the GAMESS code [63].

A detailed review of this area has been produced by Bowler and Miyazaki [64].

2.2.1 ONETEP

In the ONETEP code, the density matrix is expressed in terms of a set of spatially localised atom centred non-orthogonal basis functions ϕ_α (called non-orthogonal generalised Wannier functions - NGWFs) and a density kernel $K^{\alpha\beta}$[65, 66]:

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_{\alpha\beta} K^{\alpha\beta} \phi_\alpha(\mathbf{r}) \phi_\beta(\mathbf{r}') \tag{2.45}$$

Here, an explicit summation has been included, but going forward the Einstein convention of implicit summation of repeated indices will be used. The density matrix of an insulating material decays exponentially with separation:

$$\rho(\mathbf{r}, \mathbf{r}') \sim e^{-\lambda|\mathbf{r}-\mathbf{r}'|} \quad |\mathbf{r} - \mathbf{r}'| \to \infty \tag{2.46}$$

This means that it is possible to impose a truncation on the density matrix:

$$K^{\alpha\beta} = 0 \quad |\mathbf{R}_\alpha - \mathbf{R}_\beta| > r_{\text{cut}} \tag{2.47}$$

Where $\mathbf{R}_\alpha$ and $\mathbf{R}_\beta$ are the centres of the NGWFs labelled by α and β respectively. This truncation gives a density kernel with a number of non-zero elements which is linear in the system size. Use of efficient sparse matrix algebra algorithms permits computation of the total energy with only linear effort. This is because both the density matrix and the Hamiltonian matrix have $\mathcal{O}(N)$ elements and the trace of the product of these two matrices is required to compute the total energy.

In ONETEP, both the kernel and NGWFs are optimised to minimise the total energy of the system, subject to constraints. The procedure is carried out in a 'double loop'. The outer loop consists of the optimisation of NGWFs using a conjugate gradient method and the inner loop the optimisation of the density kernel using the LNV algorithm, which is discussed in Section 2.2.3. Figure 2.3 shows a flow diagram to illustrate the process.

One important caveat with EELS simulation in ONETEP is that the KS eigenfunctions are needed; this requires that a single $O(N^3)$ diagonalisation

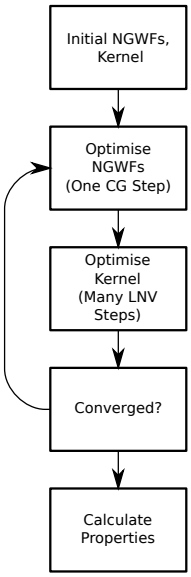


Figure 2.3: A simple flow diagram of the ONETEP double loop optimisation process. At each NGWF optimisation step the density kernel is optimised to minimise the energy.

be performed. For all but the largest of systems, the time required for this step will be negligible in comparison to that needed for the electronic minimisation.

2.2.2 NGWFs

In Section 2.1.6, some of the advantages and disadvantages of atom centred orbitals were discussed. NGWFs form a useful basis as they provide access to the small basis size of atom centred bases without the transferability problems. This is because NGWFs are optimised *in situ* to provide the best possible representation of the density matrix $\rho(\mathbf{r}, \mathbf{r}')$. These local basis functions are centred on atoms, with the centre of ϕ_α being $\mathbf{r}_\alpha$.

In ONETEP, the NGWFs are themselves expressed in terms of periodic sinc functions, meaning that ultimately ONETEP uses a basis equivalent to plane waves. This allows for easy comparison of results to those obtained from plane wave codes. The spatial localisation of the NGWFs is imposed by requiring that the coefficients of all periodic sinc basis functions centred more than the NGWF cut off radius $r_{\max}$ away from the centre of the NGWF be zero.

Let us examine the representation in more detail as described in the work of Mostofi *et al* [67]. The simulation cell is covered by a real space grid, whose points are indexed by K, L, M. There are $N_i = 2J_i + 2$ grid points along the direction of the i^{th} lattice vector. Reciprocal space lattice vectors are denoted B_i. At each grid point ($\mathbf{r}_{KLM}$) a psinc basis function $D_{KLM}(\mathbf{r})$ is placed:

$$D_{KLM}(\mathbf{r}) = \frac{1}{N_1 N_2 N_3} \sum_{p=-J_1}^{J_1+1} \sum_{q=-J_2}^{J_2+1} \sum_{s=-J_3}^{J_3+1} \exp\{\mathrm{i}(p\mathbf{B}_2 + q\mathbf{B}_2 + s\mathbf{B}_3) \cdot (\mathbf{r} - \mathbf{r}_{KLM})\} \tag{2.48}$$

The number of basis functions is thus controlled by a single parameter - the spacing of the psinc grid. This is typically expressed in terms of a cut off

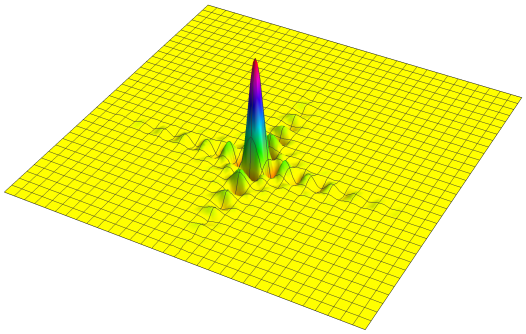


Figure 2.4: A psinc function on a 30 × 30 grid, positioned at grid point (15,12).

energy, defined in an analogous way the that used in conventional plane wave codes.

A schematic representation of a psinc function on a 2D grid is shown in Figure 2.4. Note the psinc function is mostly localised but does have tails extending throughout the simulation cell. This arises from the fact the psinc functions are, in effect, band width limited delta functions, so ringing type artefacts are unsurprising.

An NGWF is then written in terms of the psinc basis as:

$$\phi_\alpha(\mathbf{r}) = \sum_{KLM} C_{KLM,\alpha} D_{KLM}(\mathbf{r})\,, \tag{2.49}$$

With the condition that:

$$|\mathbf{r}_{KLM} - \mathbf{r}_\alpha| > r_{\max} \implies C_{KLM,\alpha} = 0\,. \tag{2.50}$$

Overlap

The NGWFs are used as a basis set to express, amongst other quantities, the density matrix and the Hamiltonian. In the case of an orthogonal basis

set, most expressions involving quantities expressed in terms of the basis set are simplified by the fact that the overlap between basis functions is the identity. This is not the case with NGWFs. As a result, an important quantity in ONETEP calculations is the NGWF overlap matrix

$$S_{\alpha\beta} = \langle \phi_\alpha | \phi_\beta \rangle \,. \tag{2.51}$$

In the course of calculations, not only $S_{\alpha\beta}$ but its inverse is required. Typical matrix inversion has $\mathcal{O}(N^3)$ cost. To avoid this, ONETEP uses an iterative inversion scheme by Hotelling [68]. This scheme has the additional advantage that during NGWF optimisation the inverse overlap matrix from the previous iteration makes a good starting point for calculation of the inverse overlap for the current iteration.

Optimisation

NGWFs are optimised via a conjugate gradients process to minimise the energy of the system at fixed density kernel. Each individual NGWF therefore adapts into an optimal form, helping ONETEP to operate with the smallest possible number of NGWFs. Because quantities in ONETEP are expressed in terms of NGWFs this in turn helps keep the cost of matrix multiplication low as the rank of the matrices involved is as small as possible. During the optimisation procedure it is possible for the optimal search direction to lead to non zero contributions to an NGWF from psincs outside its localisation radius. Care is taken to ensure these contributions are removed.

While scaling of simulation time with number of atoms in ONETEP is $\mathcal{O}(N)$, the same does not apply to scaling with other simulation parameters. Notably even small increases in NGWF radius can lead to considerable increases in simulation time.

2.2.3 Density Kernel

The density kernel underpins the ONETEP method. As it expresses the density kernel in the basis of NGWFs, it must exhibit properties linked to those of the density matrix, namely idempotency and normalisation. The first of these properties stems from the requirement that $\rho^2 = \rho$ and the second that $\mathrm{Tr}[\rho] = N_\mathrm{e}$. When expressed in terms of the density kernel these become:

$$\mathbf{KSK} = \mathbf{K} \tag{2.52}$$

$$2\mathrm{Tr}[\mathbf{KS}] = N_\mathrm{e} \tag{2.53}$$

As the energy of the system is minimised with respect to the density kernel, great care is taken to ensure Eq. 2.52 and Eq. 2.53 are maintained [69]. To ensure numerical stability, ONETEP uses a combination of multiple techniques. Initially, the canonical purification of Palser and Manolopoulos [70] generates an initial kernel. Optionally, a penalty functional method is then applied [71]. Finally, as the inner loop in the overall ONETEP minimisation process, the algorithm of Li, Nunes and Vanderbilt [72, 73] is used, with purification interspersed if needed.

Optimisation

Some of the techniques used in the optimisation of the density kernel will now be elaborated upon. This will start with a broad discussion the mathematical background before dealing with some ONETEP specific details.

An idempotent matrix has eigenvalues of either 0 or 1. Based on this observation, iterative *Purification* methods have been developed which take an initial matrix, whose eigenvalues are not exactly 1 or 0 and produces successive new matrices whose eigenvalues are closer to those values. A typical example is the McWeeny transform [74]:

$$\rho_{i+1} = 3\rho_i^2 - 2\rho_i^3 \,. \tag{2.54}$$

It is clear that an eigenvector of ρ_i, with eigenvalue λ, will also be an eigenvector of ρ_{i+1} with eigenvalue $3\lambda^2 - 2\lambda^3$. Plotting the new eigenvalue against the old, we obtain Figure 2.5a from which it is clear that eigenvalues in the range $[\frac{1}{2}\left(1-\sqrt{3}\right), \frac{1}{2}\left(\sqrt{3}+1\right)]$ are moved closer to 1 or 0. The iterative application of this method is shown in Figure 2.5b.

The *Penalty Functional* addresses the idempotency constraint by modifying the energy functional to be optimised. A new quantity, [54, 69]

$$Q[\rho] = E[\rho] + \alpha P[\rho]\,, \tag{2.55}$$

where $E[\rho]$ is the total energy, α sets the strength of the applied penalty and $P[\rho]$ is the penalty functional itself is minimised. The idea is that $P[\rho] = 0$ when ρ is idempotent and increases as ρ moves away from idempotency. Thus for a sufficiently large α, the minimum of Q will be the minimum energy $E[\rho]$ with ρ obeying the idempotency condition.

Several forms for $P[\rho]$ have been proposed. Some forms, such as that of Kohn [54] have attractive properties in principle but are practically intractable due to non-analytic behaviour [75].

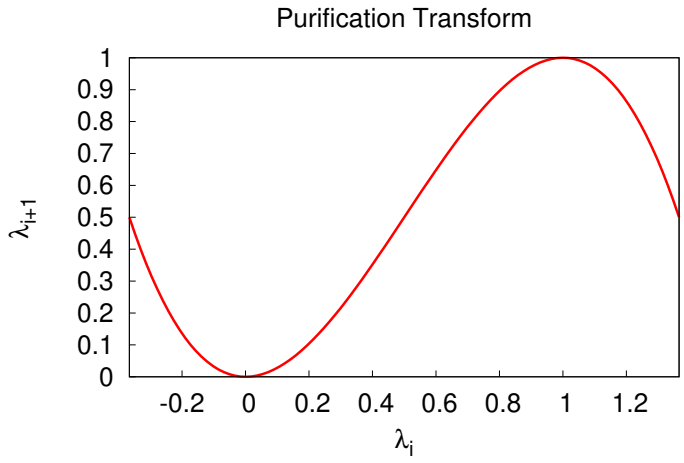
As shown in Figure 2.5b if the initial eigenvalues of a matrix lie in a suitable range, then the eigenvalues of the matrix generated by applying Eq. 2.54 once will lie in the range $[1, 0]$. This is the idea behind the LNV method. The density kernel is expressed in terms of an auxiliary kernel $\mathbf{L}$ via

$$\mathbf{K} = 3\mathbf{L}^2 - 2\mathbf{L}^3\,, \tag{2.56}$$

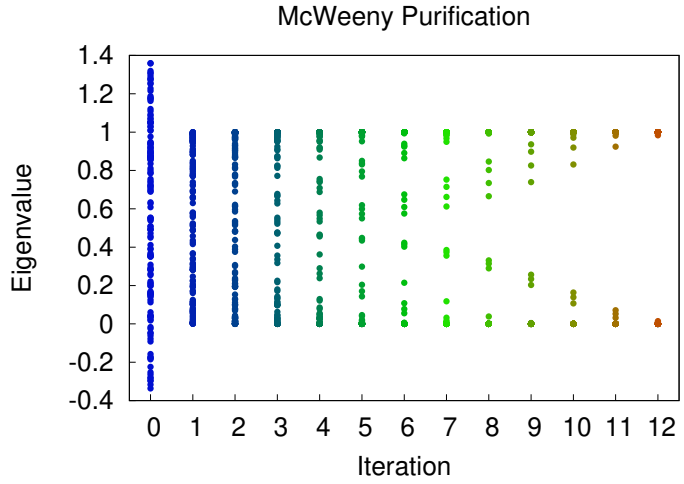
or in a non orthogonal basis such as the one used by ONETEP

$$\mathbf{K} = 3\mathbf{LSL} - 2\mathbf{LSLSL}\,. \tag{2.57}$$

The energy of the system is minimised with respect to the components of L using a conjugate gradients scheme. To ensure numerical stability, the least and greatest eigenvalue of L are checked to make sure both are within the interval $[\frac{1}{2}\left(1-\sqrt{3}\right), \frac{1}{2}\left(\sqrt{3}+1\right)]$.



(a) The purification function's effects on a single eigenvalue.



(b) The evolution of the eigenvalues of 150×150 symmetric matrix under successive applications of Eq. 2.54. The initial eigenvalues were randomly distributed in the range $[\frac{1}{2}\left(1-\sqrt{3}\right), \frac{1}{2}\left(\sqrt{3}+1\right)]$. One application of the transform is sufficient to bring all eigenvalues into the interval $[0, 1]$, this informs the LNV method.

Figure 2.5: Details of the purification transform.

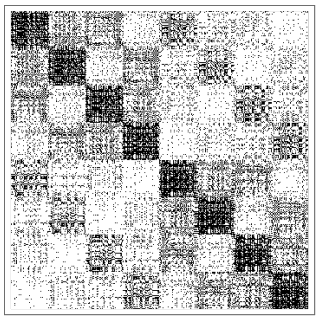


Figure 2.6: A depiction of a ONETEP density kernel. Black points indicate non-zero elements. Note the increasing sparsity away from the diagonal.

Practical Considerations

The ONETEP density kernel is represented as a sparse matrix. Internal to the code, atoms are re-ordered according to a space filling curve [76] which ensures atoms with a small spatial separation have similar indices. This has the effect of clustering the largest elements of the density kernel close to the diagonal, as shown in Figure 2.6. ONETEP exploits a hybrid MPI-OpenMP parallel scheme to efficiently manipulate large sparse matrices. This enables calculations to take advantage of modern massively parallel super computers [77].

2.2.4 FFT Boxes

In order to compute Hamiltonian matrix elements in linear time, ONETEP makes use of *FFT boxes* [78]. These are sub-regions of the simulation cell centred on each NGWF and are sufficiently large to contain all NGWFs which overlap with it. Using FFT boxes ensures that, asymptotically, the effort to compute matrix elements between NGWFs is independent of the size of the simulation cell as all matrix elements between NGWFs are computed in the fixed size FFT box rather than the whole cell. This strength only becomes apparent when the cell is large in comparison to the NGWF radius:

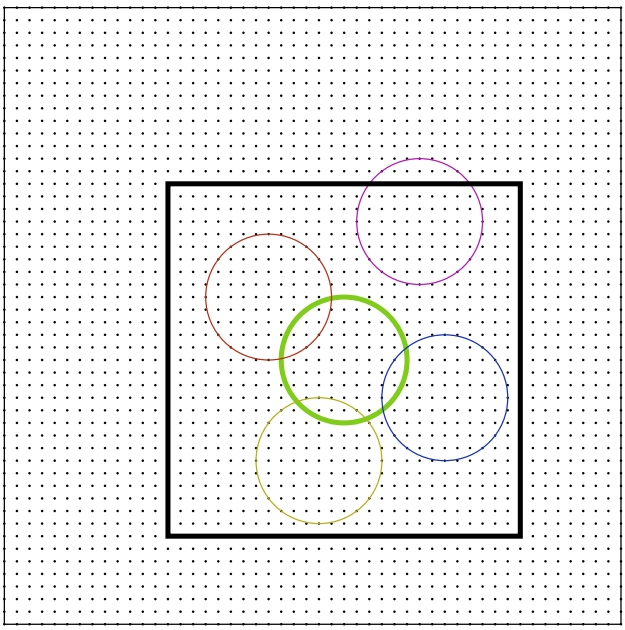


Figure 2.7: A graphical depiction of an FFT box. The whole simulation cell is outlined, with the grid points depicted also. The bold black box is the FFT box for the NGWF in bold green. Note that the three NGWFs which overlap with the central NGWF all entirely fit inside the FFT box. The Magenta NGWF, does not overlap with the central NGWF so the FFT box does not have to be large enough to contain it.

if a cell dimension is less than four times the NGWF radius, then the FFT box and cell will have the same size in that dimension. This is particularly advantageous in the case of lower dimensional nano-structures as vacuum can be included in the simulation with a reduced computational penalty compared to conventional plane wave codes [79].

2.2.5 Conduction Optimisation

While the ONETEP double loop produces NGWFs which represent the valence band well, attempts to represent conduction states using the same NGWFs

can lead to poor results. Given EEL transitions are to the conduction band states, it is essential that they are accurately represented. To overcome this problem ONETEP optimises a second set of NGWFs to represent low lying conduction band states:

$$n_c(r, r') = \chi_\alpha(r) K_c^{\alpha\beta} \chi_\beta(r') \tag{2.58}$$

This is accomplished by constructing a modified conduction Hamiltonian, where the density matrix is used to project out valence states and shift them to a higher energy:

$$H_{\text{cond}} = H - \rho(H - \sigma)\rho\,. \tag{2.59}$$

The energy of a system with this Hamiltonian is then modified, effectively optimising the lowest N_c conduction states. When choosing the number of conduction states to optimise, care must be taken to ensure that no partial degenerate manifolds are included, as this can cause the optimisation to fail to converge. As conduction states can be rather more delocalised than valence states, it is sometimes necessary to increase the radius of conduction NGWFs above that used for valence NGWFs. An automated method for optimising conduction states within an energy window has recently been added to ONETEP, but for this work the control provided by directly specifying the number of states was preferred.

The optimisation of conduction NGWFs is not self consistent and is consequently much less computationally intensive than a ground state calculation. The modified Hamiltonian Eq. 2.59 introduces new terms into the gradient of energy with respect to NGWFs and Kernel, but beyond this enables existing ONETEP internal routines to be reused.

The conduction optimisation method was originally tested in isolated organic molecules [80], but in the course of this work has been found to work well in bulk systems too (see Chapter 3).

Once a set of conduction NGWFs has been produced it is possible, if desired, to diagonalise the system Hamiltonian expressed in a basis of both

valence and conduction NGWFs. The *joint basis* allows for both occupied and low lying conduction states to be well represented. In subsequent chapters, when reference is made to expressing conduction band wavefunctions, this should be understood to refer to expressions in terms of the joint basis, unless explicitly indicated otherwise. Following diagonalisation, wavefunctions may be expressed in terms of NGWFs using a matrix notation:

$$\psi_i = M_i^{\alpha}\phi_{\alpha}\,. \tag{2.60}$$

The M matrix defined above provides a means of converting between NGWF and wavefunction representations, and is crucial in the computation of EELS spectra.

PAW in ONETEP

As described in Section 2.1.5, PAW involves the KS eigenstates. Clearly this formulation cannot be used as is in ONETEP, so modifications must be made in order to formulate PAW in terms of the density matrix and other quantities which can be computed in $\mathcal{O}(N)$ time. Hine shows [81] that the all-electron density matrix can be expressed in the PAW formalism as

$$\rho = \widetilde{\rho} + \sum_{\nu\mu}(|\varphi_{\nu}\rangle\langle\varphi_{\mu}| - |\widetilde{\varphi_{\nu}}\rangle\langle\widetilde{\varphi_{\mu}}|)\langle\widetilde{p}^{\nu}|\widetilde{\rho}|\widetilde{p}^{\mu}\rangle\,. \tag{2.61}$$

The form of the idempotency and normalisation constraints (Eqs. 2.52, 2.53) can be retained if the NGWF overlap matrix is modified to take into account PAW terms:

$$S_{\alpha\beta} = \langle\phi_{\alpha}|[1 + |\widetilde{p}^{\nu}\rangle(\langle\varphi_{\nu}|\varphi_{\mu}\rangle - \langle\widetilde{\varphi}_{\nu}|\widetilde{\varphi}_{\mu}\rangle)\langle\widetilde{p}^{\mu}|]\phi_{\beta}\rangle\,. \tag{2.62}$$

This redefinition of the overlap matrix enables most of ONETEP's internal code to be retained for PAW calculations without modification.

The projector augmented wave method is agnostic to the basis set used, however the ONETEP plane wave basis requires some particular consideration.

Here we will discuss the calculation of PAW corrections to matrix elements of an arbitrary operator $\hat{O}$ between NGWFs, ϕ_α and an arbitrary projector function b. The projectors and partial waves of the PAW method are most naturally stored on radial logarithmic grids, which permit the rapidly varying functions to be treated with great accuracy. Conversely, the NGWFs exist on a Cartesian grid, so in order to perform NGWF-projector matrix element integrals, a transformation of the projectors onto the Cartesian grid is necessary:

$$\langle\phi_\alpha|\hat{O}|b\rangle = \underbrace{\langle\phi_\alpha|\hat{O}|b\rangle}_{\text{Cartesian Grid}} + \sum_i \langle\phi_\alpha|\widetilde{p}_\mathrm{i}\rangle \underbrace{\left(\langle\varphi_\mathrm{i}|\hat{O}|b\rangle - \langle\widetilde{\varphi}_\mathrm{i}|\hat{O}|b\rangle\right)}_{\text{Radial Grids}} \tag{2.63}$$

The calculation of the grid part of the matrix element is novel in this work and will be discussed separately in Chapter 3. The radial grid part of the code was pre-existing [81], and will be discussed here.

The process for calculating the radial parts of the correction term is quite simple:

- In advance, radial integrals of the core wavefunctions with the partial waves are computed.
- For each atom, a double loop over all core wavefunctions and partial waves is undertaken.
- The previously computed radial integrals are weighted with an appropriate real gaunt coefficient and placed in a result array. This step exploits the fact that the position operator can be expressed in terms of the real spherical harmonics.

With this theoretical basis in place, it is now possible to go forward and develop a method for computing EEL spectra in ONETEP.

2.3 Conclusion

The methods discussed in this chapter have revolutionised electronic structure calculation. Many previously impossible calculations are now a matter of routine. Ever larger and more complex systems are being studied, both as a result of improvements in computer speed and cost and, crucially, due to the development of methods which can exploit those improvements. In particular, the ONETEP code has been applied to a great variety of systems including dyes in solution [82], metallic nano particles [83], numerous biological systems [84], carbon nano structures [85] and 2D materials [86]. Systems with 21000 atoms have been simulated [77].

Chapter 3

Implementation of EELS in LS-DFT

The work underlying this chapter has been published in *Journal of Physics: Condensed Matter* as "Simulation of electron energy loss spectra of nanomaterials with linear-scaling density functional theory" DOI: 10.1088/0953-8984/28/19/195202 [87].

The key objective of this work is to demonstrate the implementation and utility of an EELS simulation framework within the ONETEP code. This chapter will focus on implementation and initial testing. The first section of this chapter will address the theoretical basis for EELS simulation from ground state DFT calculations. Discussion will then move to the specific details of implementation in ONETEP, with focussing on the details of how the projector augmented wave method is used. Comparison to existing DFT codes, both all-electron and plane wave pseudopotential, is made. Finally, calculated spectra are compared to experimental data for a variety of test systems.

3.1 Theoretical Basis

3.1.1 Fermi's Golden Rule

The objective of EELS simulation is to compute the loss function of a system

$$L = \Im\{\frac{1}{\epsilon}\}\,, \tag{3.1}$$

Where ϵ is the complex dielectric function

$$\epsilon = \epsilon_1 + \mathrm{i}\epsilon_2\,. \tag{3.2}$$

Both ϵ_1 and ϵ_2 are frequency dependant. For high frequencies like those associated with core loss EEL transitions, ϵ_1 tends to 1 and ϵ_2 tends to a small value. Using these limits Eq. 3.1 may be written

$$\begin{aligned} L &= \Im\{\frac{1}{\epsilon}\} \\ &= \Im\{\frac{1}{\epsilon_1 + \mathrm{i}\epsilon_2}\} \\ &= \Im\{\frac{\epsilon_1 - \mathrm{i}\epsilon_2}{\epsilon_1^2 + \epsilon_2^2}\} \\ &\approx \Im\{\epsilon_1 - \mathrm{i}\epsilon_2\} \\ &\approx \epsilon_2\,. \end{aligned} \tag{3.3}$$

The problem of computing ϵ_2 may be tackled using Fermi's Golden Rule. To apply the rule, the passing electron beam is treated as a time varying perturbation with arbitrary (but small enough for time dependent perturbation theory to be a valid approximation) amplitude A, angular frequency ω and momentum transfer $\mathbf{q}$:

$$V(\mathbf{r}, t) = A\exp\{\mathrm{i}(\omega t - \mathbf{r}\cdot\mathbf{q})\}\,. \tag{3.4}$$

Substituting this into the textbook expression gives:

$$\epsilon_2(\omega) = \frac{1}{\Omega}\sum_c\sum_i |\langle\psi_i|\exp(\mathrm{i}\mathbf{r}\cdot\mathbf{q})|\psi_c\rangle|^2\delta(E_i - E_c - \omega)\,, \tag{3.5}$$

where the ψ_i are single particle slater determinant orbitals. Here, the c sum runs over all core wavefunctions, and the i sum runs over all unoccupied orbitals. Note that this limits the region of the spectrum computed to the core loss regime. The delta function conserves energy. Atomic units have been used. The exponential term represents the perturbing spatial potential of the electron beam. Adopting the dipole approximation (assuming that the magnitude of the momentum transfer, $\mathbf{q}$, is small) enables the term to be written as

$$\exp(\mathrm{i}\mathbf{r}.\mathbf{q}) \approx 1 + \mathrm{i}\mathbf{r}\cdot\mathbf{q}\,. \tag{3.6}$$

This yields the matrix element:

$$\begin{aligned}\langle\psi_i|\exp(\mathrm{i}\mathbf{r}.\mathbf{q})|\psi_c\rangle &\approx \langle\psi_i|1+\mathrm{i}\mathbf{r}\cdot\mathbf{q}|\psi_c\rangle \\ &\approx \mathrm{i}\mathbf{q}\cdot\langle\psi_i|\mathbf{r}|\psi_c\rangle\,,\end{aligned} \tag{3.7}$$

where in the final step orthogonality of core and conduction states was used. The final expression for the imaginary part of the permittivity is

$$\epsilon_2(\omega) = \frac{1}{\Omega}\sum_c\sum_i |\mathbf{q}\cdot\langle\psi_i|\mathbf{r}|\psi_c\rangle|^2\delta(E_i - E_c - \omega)\,. \tag{3.8}$$

Writing ϵ_2 using the dipole approximation has the advantage that a single DFT run can provide the matrix elements required to compute an EEL spectrum for a material at a range of different (small) momentum transfers.

Here, an important approximation has been made which must be emphasised. It has been assumed that the Kohn-Sham conduction band eigenstates are a good representation, in terms of energy and morphology, of the true excited states of the system. This may appear to be a rather crude approximation, but it has been shown that in materials which are not strongly correlated, low energy KS conduction states match the behaviour of low energy quasiparticle excitations well, with the caveat that the energy gap obtained by KS DFT is typically several tens of percent less than the experimental value. The adequacy of using KS conduction states as an approximation to the true states is borne out by the success of TD-DFT and DFT ARPES

simulations. The approximation is also needed to make progress. Its validity is demonstrated later in this chapter by comparisons to experiment.

3.1.2 Existing methods for EELS Calculation in Plane Wave Codes

Band structure methods for EELS calculation are now over thirty years old. Early examples of the use of the method can be found in the work of Linder *et al.*, who studied the venerable test system MgO [88]. This early work was based on the Korringa–Kohn–Rostoker method [89, 90].

The increasing power of electronic structure codes lead to their adoption for the calculation of edge structure. Both augmented plane wave and full plane wave codes such as CASTEP were equipped with this functionality [25].

At a similar time, workers were also focussing on methods to calculate x-ray absorption near edge spectra. This class of spectrum is modelled using a method that is effectively identical to that used for EELS. Work by Taillefumier *et al.* discussed the use of the projector augmented wave [91] method in this case.

The method used to simulate EELS in CASTEP is outlined in [92] and is similar to the method described here, save the modifications needed to work in an NGWF basis. The CASTEP implementation has been used to study a variety of systems including nuclear fuels [93], magnesium oxide materials of use in tunnel junctions [94], the impact of adsorbates on the EELS spectra of platinum catalysis [95] and the study of organic molecules reacting an oxide surfaces [96].

Improvements to the technique have tended to focus on better treatment of core hole effects. Mauri demonstrated the use of modified pseudopotentials to account for core excitonic interactions in diamond [97]. Gao *et al.* demonstrated an improvement in experimental agreement, seen when core holes are included in simulations of simple compounds of first row elements [98].

Advances in the CASTEP code have put small nano structures within computational reach [99]. The CASTEP *on the fly* pseudopotential generator [100] enabled simplified generation of core hole data sets for CASTEP EELS calculations.

Similar work on core hole effects has been undertaken using the VASP code [101]. Recent work using VASP has included the development of a hybrid approach, where the local density of states is computed using the VASP code and the transition matrix elements are computed for a smaller system using WIEN2K [102].

3.2 Computing Matrix Elements

We now deal with the details of the computation of EEL spectra in ONETEP. As discussed in Chapter 2, there are two components to a PAW corrected matrix element. The first is formed from the PAW correction and is computed on radial grids. The second is computed on the regular Cartesian grid. The degree of contribution from the two parts depends on the operator whose matrix element is being computed ($\mathbf{r}$ in this case) and the projector ket forming the left hand side of the element (a core wavefunction ψ_c in this case). Low Z elements such as carbon will see a greater relative contribution from the Cartesian grid. In all cases however, the best accuracy can only be achieved if both parts are calculated correctly. To this end, it is essential to form the most accurate possible representation (for a given cutoff energy) of the $\mathbf{r}|\psi_c\rangle$ ket. There are two possible methods for computing this. One is to directly take the product of the position operator with a core wavefunction on a Cartesian grid. This route suffers from a major flaw in that it results in a very poor representation of the $\mathbf{r}|\psi_c\rangle$ product - multiplying the two frequency limited real space representations leads to the omission of contributions to the product from the cancelling interactions of high frequency components which would otherwise be included. Figure 3.1 provides a 1D example of this

problem. In the limit of an infinite basis set, this method would reproduce the correct result. Practical bases can not provide the required accuracy of representation.

A superior method is to take the gradient of the core wavefunction in Fourier space and transform the result back to real space. This exploits the Fourier representation of the position operator:

$$\sum_{\mathbf{G}}^{\mathbf{G}_{\max}} (\nabla_{\mathbf{G}} \psi_{\mathrm{c}}(\mathbf{G})) \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r}) = -\mathrm{i}\mathbf{r}\psi_{\mathrm{c}}(\mathbf{r})\,. \tag{3.9}$$

This result may be proven by noting:

$$\begin{aligned} f(\mathbf{G}) &= (\nabla_{\mathbf{G}} \psi_{\mathrm{c}}(\mathbf{G})) \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r}) \\ &= \nabla_{\mathbf{G}}(\psi_{\mathrm{c}}(\mathbf{G}) \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r})) - \psi_{\mathrm{c}}(\mathbf{G})\nabla_{\mathbf{G}} \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r}) \\ &= \nabla_{\mathbf{G}}(\psi_{\mathrm{c}}(\mathbf{G}) \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r})) - \psi_{\mathrm{c}}(\mathbf{G})\mathrm{i}\mathbf{r} \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r}) \end{aligned} \tag{3.10}$$

by the product rule. Inserting Eq. 3.10 into a sum over $\mathbf{G}$ vectors we obtain

$$\begin{aligned} &\sum_{\mathbf{G}}^{\mathbf{G}_{\max}} (\nabla_{\mathbf{G}}(\psi_{\mathrm{c}}(\mathbf{G}) \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r})) - \psi_{\mathrm{c}}(\mathbf{G})\mathrm{i}\mathbf{r} \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r})) = \\ &\sum_{\mathbf{G}}^{\mathbf{G}_{\max}} \nabla_{\mathbf{G}}(\psi_{\mathrm{c}}(\mathbf{G}) \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r})) - \mathrm{i}\mathbf{r} \sum_{\mathbf{G}}^{\mathbf{G}_{\max}} \psi_{\mathrm{c}}(\mathbf{G}) \exp(\mathrm{i}\mathbf{G}\cdot\mathbf{r}) = \\ &\qquad -\mathrm{i}\mathbf{r}\psi_{\mathrm{c}}(\mathbf{r}) \end{aligned} \tag{3.11}$$

where in going from the first to the second line, the distributive nature of addition and the fact that the sum does not depend on $\mathbf{r}$ have been used. Going from the second to third line relies on the observation that the first term on line two will vanish due to periodic boundary conditions, and that the second term is a Fourier series for $\psi_{\mathrm{c}}(\mathbf{r})$.

3.2.1 NGWFs and FFT Boxes

Internally, ONETEP makes extensive use of FFT boxes. FFT boxes are sub grids of the whole simulation cell, large enough to accommodate an NGWF

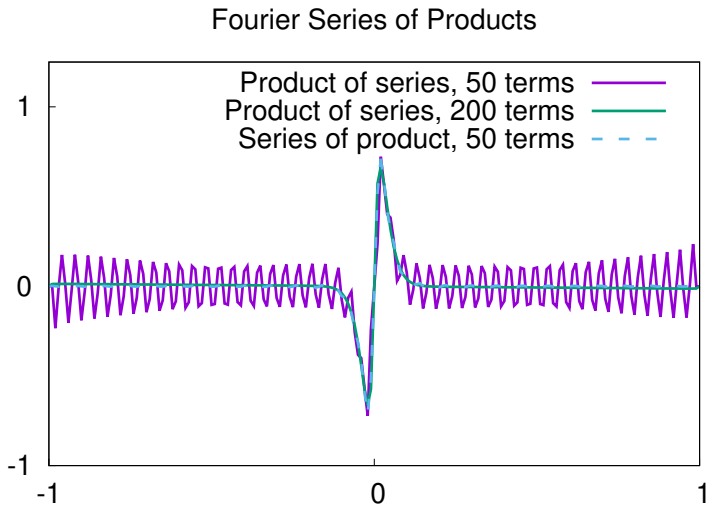


Figure 3.1: Taking the product of two band width limited Fourier series in real space may lead to a completely inaccurate representation of the product of the two original functions. This figure depicts an attempt to represent the periodically repeated function $100x\exp(-55|x|)$ as the product of the Fourier series for x and $100\exp(-55|x|)$. Note that while the series of the product (equivalent to the Fourier space gradient method) produces a well converged result when 50 terms are included in its Fourier series, around 200 terms are required to suppress the ringing phenomenon seen in the product of series method.

and all NGWFs which overlap with it. To make use of existing code, the Cartesian part of the PAW matrix element is computed in an FFT box. The procedure for this is based on the representation of the position operator in Fourier space. In reality, ONETEP acts on *batches* of projectors (core wavefunctions) in a single operation. For clarity a single core wavefunction is used here:

1. Take $\nabla_{\mathbf{G}}|\psi_c(\mathbf{G})\rangle$ - in onetep the result is deposited in an FFT box
2. Exchange real and imaginary FFT box grids and multiply the new real grid by -1 (This removes the factor of $-$i from Eq. 3.9)
3. FFT to real space
4. The FFT box now contains $\mathbf{r}|\psi_c(\mathbf{r})\rangle$
5. Compute $\langle\phi_\alpha|\mathbf{r}|\psi_c(\mathbf{r})\rangle$ for all NGWFs which have non zero coefficients in the FFT box
6. The resulting matrix elements are converted into elements involving the conduction band wave functions using the result $\langle\tilde{\psi}_i| = M_i^{\dagger\alpha}\langle\phi_\alpha|$

3.3 Matrix Elements to Spectra

Using Eq. 3.8 to compute an EEL spectrum from matrix elements was accomplished by the OPTADOS [103, 104] tool. OPTADOS is a pre-existing tool, which was originally developed to process CASTEP output files. ONETEP was extended to output a `.elnes_bin` file in the same format as the CASTEP file, which can be read in by OPTADOS and used with a variety of broadening schemes. For all our comparisons, an isotropic average over momentum transfer was taken. It should be noted, however, that OPTADOS has the facility to specify a momentum transfer.

3.3.1 Broadening Schemes

The process of computing a finished spectrum requires the replacement of the delta function in Eq. 3.8 with a suitable broadening function. Lorentzian (representing lifetime broadening effects) and Gaussian (finite energy resolution and other effects) are commonly chosen. The characteristic width of these functions can either be specified as a fixed value or determined using the optical matrix elements of the material. In the present case, fixed broadening was used.

3.4 Core Holes

When an EEL transition takes place, an electron is excited from a core orbital to the conduction band of the sample. This excitation leaves a hole in the core shell of the excited atom. The positively charged hole and negatively charged electron interact and this leads to changes in the resulting spectra. In order to accurately model experimental spectra, this interaction must be taken into account [98, 105]. One approximate method for this, which is readily implemented in plane wave / PAW codes, is to use a modified pseudopotential / PAW dataset generated for an atomic configuration where (some fraction of) a core electron has been excited to valence orbitals. Some example configurations for carbon are: $1s^2 2s^2 2p^4 \rightarrow 1s^1 2s^2 2p^5$ or $1s^2 2s^2 2p^4 \rightarrow 1s^{1.5} 2s^2 2p^{4.5}$. A second, simpler approach is used by some codes. This is the 'Z+1' method. This technique simply replaces the pseudopotential of the atom being investigated with that of the atom with one greater atomic number.

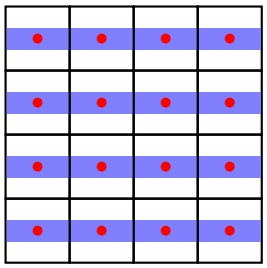


Figure 3.2: The supercell approximation in 2D for a slab. A defect (red dot) is enclosed in a simulation cell which is repeated to cover all space. The simulation cell consists of a slab of material (blue) surrounded by vacuum (white). As the simulation cell is made larger, the limit of a single isolated defect would eventually be reached. However, leading terms in the electrostatic and elastic interactions decay slowly, as $\frac{1}{r}$ with increasing separation r. Using only a finite size simulation cell introduces errors. Coping with these is the subject of Chapter 6

3.5 Supercells

When modelling defects in materials which, in principle, have infinite extent in one or more dimensions, a variety of efficient approximations may be needed. Perfect bulk materials are treated using periodic boundary conditions, as described in Section 2.1.6. The introduction of a defect (be that an impurity, a vacancy or even a surface) disrupts the perfect translational symmetry which underpins the use of periodic boundary conditions. To address this issue, the *supercell approximation* is made. The defect is included in a larger simulation cell than the primitive cell of the material in the hope that, if the cell is sufficiently large, the defect will behave as though it is isolated in the bulk. This idea is correct in the limit that the supercell is made arbitrarily large, but practical considerations mean that at some finitely sized supercell must be used and the problem of spurious interactions addressed. Figure 3.2 provides a 2D depiction of the supercell approximation.

3.6 k point sampling

Often, a fine k-point grid is required in order to produce well converged spectra. In the case of ONETEP, only the Γ point is available. To overcome this problem large simulation cells are used, which effectively reproduce the effects of k point sampling. For example, a 5×5×5 super cell produces results equivalent to a 5×5×5 k-point grid on the original unit cell. While this approach may seem inefficient, it has one great advantage: in systems being simulated using a core hole, the different pseudo potential effectively creates a defect in the material. In a real system, these (charged!) defects would occur in isolation. In the simulation with periodic boundary conditions this is not the case. Using large simulation cells ensures minimal interaction between the core hole atoms and their periodic images.

3.7 Edge Offsets

As described in the introduction, specific elements generate energy losses with characteristic onset energies. It would be useful to be able to estimate these in simulation. It is worth noting here the limitations of the approach used: Removing core states from the calculation through the use of pseudopotentials means that the energy of the initial electron level in an EELS transition is not available. Furthermore, DFT is well known to underestimate band gap energies by significant amounts, thus the edge onsets computed will not be directly comparable to experiment. Instead, the objective is to produce an edge onset which may be compared between simulations, particularly two simulations with identical cells but where spectra for several different atoms of the same species are being computed. This permits the resulting spectra to be plotted on a common energy axis. Also, given spetra for multiple atoms in the same cell share an energy axis it is possible to readily compute space averaged EEL spectra.

Here, the method of Mizoguchi *et al.* [106] is adopted. Three ingredients are used to calculate the resulting edge offset:

1. The energy difference between a whole cell simulation in the ground state, and a whole cell simulation with a core hole. In the latter case, the core hole electron is placed into the lowest unoccupied orbital of the system.

2. The energy difference between an all-electron simulation of the isolated atom in its ground state, and the same atom with a core electron promoted.

3. The energy difference between the pseudo atom in its ground state, and the same pseudo atom with a core hole potential and promoted electron.

This method does require extra simulations, however item 2 can be obtained from the pseudopotential generator used and item 3 from the ONETEP atom self consistent field solver. These calculations are fast (much less than 1 minute). Only computing item 1 requires significant computational outlay.

The Mizoguchi formula for edge energy is:

$$\begin{aligned} E_{\text{edge}} =&(E_{\text{sys+ch+e}} - E_{\text{sys,gs}}) \\ &+(E_{\text{aeatom+ch+e}} - E_{\text{aeatom,gs}}) \\ &-(E_{\text{psatom+ch+e}} - E_{\text{psatom,gs}}) \end{aligned} \tag{3.12}$$

Note the structure of Eq. 3.12 is reminiscent of the PAW correction: the contribution of the pseudised atom is subtracted off, and that of the all-electron atom is added back in. When experimental and simulated spectra are plotted on the same graph, two x axes are used: the upper axis for the simulated spectrum and the lower for the experimental spectrum.

3.8 Test Cases

To confirm that the implementation was correct, several test simple test cases were chosen. The systems were selected on the basis of their simplicity and the availability of experimental spectra for comparison. The systems are listed in Table 3.2. These systems were constructed using data from the Inorganic Crystallographic Structure Database (ICSD) [107], with the exception of the BN nanotube which was generated using a modified version of the Atomic Simulation Environment python package. Structures were not relaxed: Test simulations on a few systems indicated there was only a very slight difference in spectra when relaxation was applied and this difference would likely be imperceptible at experimental resolution.

3.8.1 Plane Wave Pseudopotential DFT

For comparison to plane wave DFT, the CASTEP code (Version 8.0) was selected [92, 100]. An academic release of ONETEP (Version 4.3.3.4) was used. Both codes were operated with an energy cut off of 800 eV and the PBE functional [37]. Different PAW datasets were used as CASTEP can only compute EEL spectra when using its *on the fly* generated data sets. The ONETEP code was equipped with the data sets of Jollet, Torrent and Holzwarth (JTH) [108]. Both these data sets have been assessed by the *Delta* project [109] and found to yield accurate results in a number of situations. No core holes were used for these comparisons. Γ point sampling was used in both codes. The ONETEP NGWF radii were set to $10a_0$ (5.3Å) for both conduction and valence calculations in all the materials studied. No density kernel truncation was used. Conduction optimisation in ONETEP was conducted by examining the density of states plot created using the valence NGWFs, and identifying a point in the conduction manifold where a gap (typically >0.1 eV) occurs. All conduction states below this point were optimised, sometimes with a few (approximately 10) extra states included for the first few iterations. More

Element	Muffin Tin Radius / Å
C	0.95
Mg	1.16
O	0.95

Table 3.1: The muffin tin radii used in all-electron calculations

details on the conduction optimisation procedure can be found in Section 2.2.5.

3.8.2 All Electron DFT

All electron DFT treats core electrons on an equal footing to valence electrons and provides a powerful test of the ability of the PAW scheme to correctly calculate matrix elements of the position operator between core and conduction states. Furthermore, the specific implementation (the ELK code [110] was used) computes EEL matrix elements of the form $\langle\psi_i|\exp(i\mathbf{r}.\mathbf{q})|\psi_c\rangle$, meaning that this comparison also provides a test of the validity of the dipole approximation. Due the computational intensity of all-electron calculations, smaller simulation cells were used and the code modified to exploit MPI parallelism.

The following computational parameters were used: for diamond, a $2 \times 2 \times 2$ supercell with $10 \times 10 \times 10$ k-point sampling; For MgO, an 8 atom cell was used along with $8 \times 8 \times 8$ k-point sampling. The LDA functional was used, with the 'Z+1' approximation for core hole effects. A basis set size parameter, `rgkmax`, of 7 was used. The muffin tin radii used are given in Table 3.1.

3.9 Convergence

For the studies in this section the computational parameters not being investigated were set to the values used earlier (energy cutoff 800eV, $10a_0$ NGWF

System	Space Group	Atoms	Ref
Diamond	$Fd\bar{3}m$	216	[111]
Gaphite	$P6_3/mmc$	200	[112]
MgO	$Fm\bar{3}m$	216	[113]
Silicon	$Fd\bar{3}m$	216	[114]
h-BN	$P6_3/mmc$	200	[115]
c-BN	$F\bar{4}3m$	216	[116]
BN nano tube	N/A	200	

Table 3.2: The systems used to test the ONETEP EELS implementation. The nanotube structure uses the lattice parameters for bulk h-BN. It was not relaxed.

radius, approximately the first 20eV of conduction band optimised).

Energy Cutoff

A series of simulations were conducted in the 216 atom MgO cell to ascertain the impact of varying the energy cut-off on the predicted spectrum. As can be seen from Figure 3.3, there is a minimal change in the predicted spectrum once the cutoff energy is greater than 600 eV. It is worth noting that for some systems a higher cutoff may be required. For example, systems with surfaces may experience unphysical negative charge densities in the vacuum above the surface even when using a cutoff value which worked well for the bulk.

NGWF radius

Another important parameter in ONETEP simulations is the NGWF radius. Given that increasing this quantity in small simulation cells is difficult (an NGWF cannot overlap with its own periodic image), these simulations were conducted in a 1728 atom MgO cell. NGWF radii of $8a_0$ and $10a_0$ were used

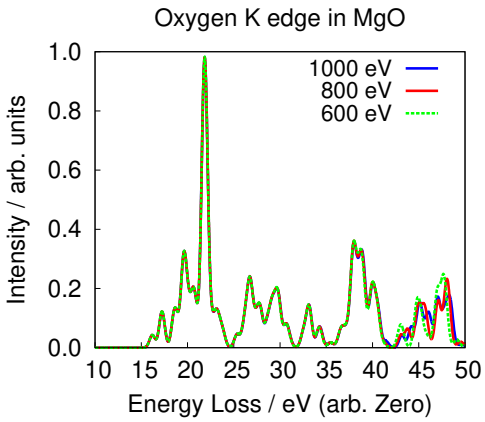


Figure 3.3: Convergence of EEL spectrum for oxygen K edge in MgO with cut-off energy in the ONETEP code. Even at 600 eV, the spectrum is well converged for energy losses of less than 40 eV. This spectrum has no broadening, to exaggerate differences. The minor changes in the > 45eV range are obscured when the spectrum is broadened with a reasonable experimental resolution.

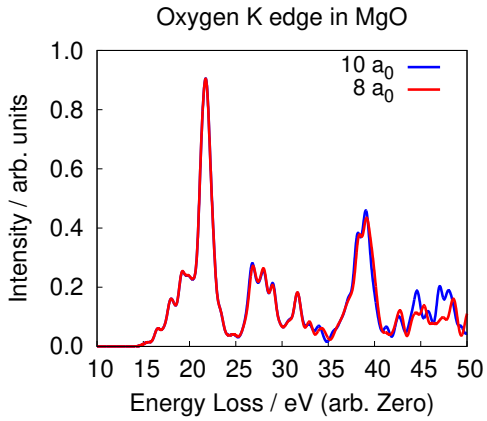


Figure 3.4: Convergence of EEL spectrum for oxygen K edge in MgO with radius of NGWF (both conduction and valence). It can clearly be seen that for the range of energies of interest for experimental comparison, an $8a_0$ cutoff radius is sufficient.

for both valence and conduction NGWFs. As can be seen from Figure 3.4, the $8a_0$ calculation is well converged for the first 30eV of energy loss.

Number of Conduction States Optimised

Another important parameter is the number of conduction states selected for optimisation. Given that the precise number of states which can be optimised is system dependent, the following strategy was adopted for this investigation: Three spectra were prepared, one using only the valence NGWFs (no conduction states), a second using conduction states optimised to represent the first 10eV of conduction band, and a third with states optimised to represent the first 20eV of the conduction band. Once again, the MgO system was used. The results are depicted in Figure 3.5 and it is clear that optimising even a few conduction states is sufficient with the 105 state case producing an effectively identical spectrum to the 692 state case over the first 30eV of energy loss.

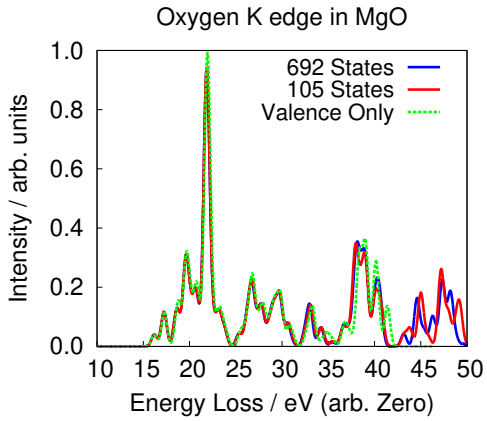


Figure 3.5: Convergence of EEL spectrum for oxygen K edge in MgO, with number of conduction states optimised. For this material, even the valence basis performs adequately. The 105 state case covers approximately 6eV above the edge onset, while the 692 state case covers about 22eV.

System size

As previously mentioned, there are two factors likely to impact the quality of a simulated spectrum as system size changes: k-point sampling, described in Section 3.6, and interactions between core holes due to periodic boundary conditions. To assess the degree of convergence at the 200 atom scale, the capacity of ONETEP to simulate thousand atom systems was used. A $2\times2\times2$ replica of the 216 atom MgO cell was constructed, containing a total of 1728 atoms. The oxygen K edge was then computed in this cell and compared to the spectrum calculated for the original. When minimal broadening is used, the result of the finer effective k-point sampling is clearly visible, with peaks being much sharper in the larger system. When an experimentally realistic broadening is applied however, this distinction is removed and the two spectra are visually almost indistinguishable (Figure 3.6). That there are no gross changes in the structure of the spectrum indicates that even at the 216 atom scale, the core hole potentials are sufficiently isolated to remove

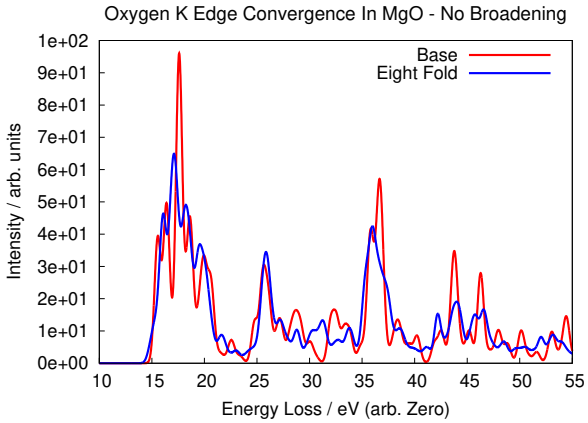


Figure 3.6: The two oxygen K edge spectra computed for MgO in two different cells.

spurious interaction.

The convergence behaviour of the Mizoguchi edge offset was also investigated. For this, an additional 64 atom cell was constructed for use with CASTEP (This cell is too small to accommodate ONETEP NGWFs). The resulting edge offset energies were found to converge quite slowly with system size. Going from a 64 atom cell ($6 \times 6 \times 6$ k-point sampling) to a 216 atom cell lead to a change of 484 meV, while going from a 216 atom cell to a 1728 atom cell led to a difference of 240 meV. Small differences in computed edge offsets are relevant when combining spectra for multiple atoms in a system. Good convergence of these quantities, especially when the absolute differences involved are small, is therefore essential.

3.10 Results

3.10.1 Comparison to Projected Density of States

The matrix elements involved in the computation of an EEL spectrum and a projected density of states are very similar. For an EEL spectrum (in the

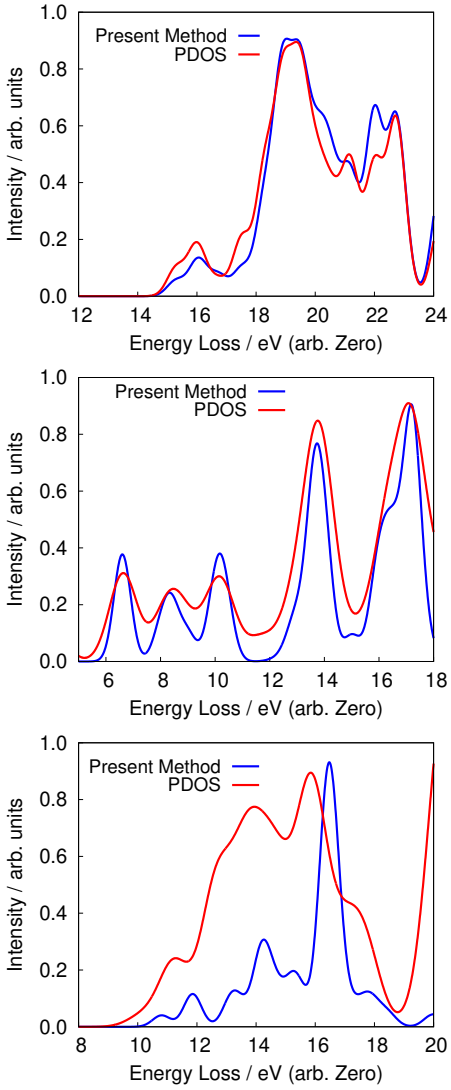


Figure 3.7: Comparisons between the projected densities of state and predicted EEL spectra for three systems. The PDOS is for the p angular momentum channel of the same atom for which the EEL spectrum was calculated. The first figure shows diamond, the second graphite and the third is for an oxygen atom in MgO.

dipole approximation) we have:

$$\langle \psi_i | \mathbf{r} | \psi_c \rangle \,. \tag{3.13}$$

Whereas for a PDOS the elements take the form

$$\langle \psi_i | \psi_{\text{atomic}} \rangle \,. \tag{3.14}$$

By noting that the product $\mathbf{r}|\psi_c\rangle$ has the same symmetry as an atomic orbital with an angular momentum one higher than ϕ_c is is clear that a good degree of qualitative agreement between an EEL spectrum and a PDOS for an angular momentum channel one higher that that of the core orbital involved in the simulated EEL transition. This is because any matrix element which vanishes by symmetry in the case of the EEL calculation will also vanish in the PDOS calculation.

As figure 3.7 shows, for simple homogeneous systems like diamond and graphite - the PDOS is a good fit with the EEL spectrum. For more complex materials such as MgO the quality of agreement diminishes.

For a further demonstration of the advantage to the present method over simply examining the PDOS for a given angular momentum channel consider the experimental EEL spectrum for cubic boron nitride, which is compared to the p ($l = 1$) angular momentum channel PDOS of a single boron atom (using a whole core hole). This is done in figure 3.8. A substantial discrepancy is clear.

3.10.2 Plane Wave DFT

Figure 3.9 shows the results of simulations for six edges (oxygen K edge in MgO, magnesium K edge in MgO, carbon K edge in graphite and diamond, and silicon K and $L_{2,3}$) for both CASTEP, a plane wave pseudopotential code, and ONETEP. This differ demonstrates an excellent agreement between ONETEP and CASTEP. In the case of K edges of light elements, the

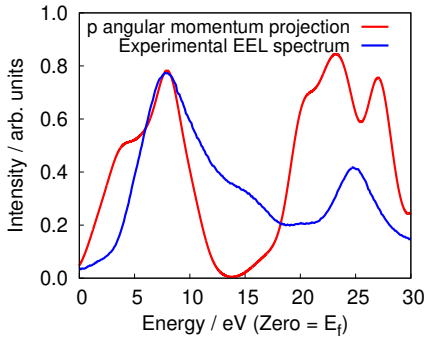


Figure 3.8: Comparison between an experimental spectrum for cubic boron nitride (from [117]) with the $l = 1$ angular momentum projection onto the site of a boron atom. A core hole was placed on the boron atom to account for interactions between the excited electron and the resulting vacancy in the core.

agreement is essentially perfect. In heavier elements (silicon in this case), slightly worse agreement is seen at higher energies. This may be due to a poorer representation of conduction states by the NGWF basis in these materials. This provides validation that matrix element generation is correct within the PAW formalism. Additionally, this comparison tests the validity of the conduction optimisation scheme in bulk solids.

3.10.3 Comparison To All Electron DFT

Comparison to all-electron DFT results is shown in Figure 3.10. Once again, a very good level of agreement is seen. In the case of MgO, there is a minor shift in the relative energy of peaks, with the second peak lying higher in energy in the ONETEP calculation than the all-electron one. This pattern is reversed for the third peak. A possible cause of this is the greatly different methods of including core hole effects. Overall, the strength of agreement provides confidence that the PAW formalism is effective in this regime and

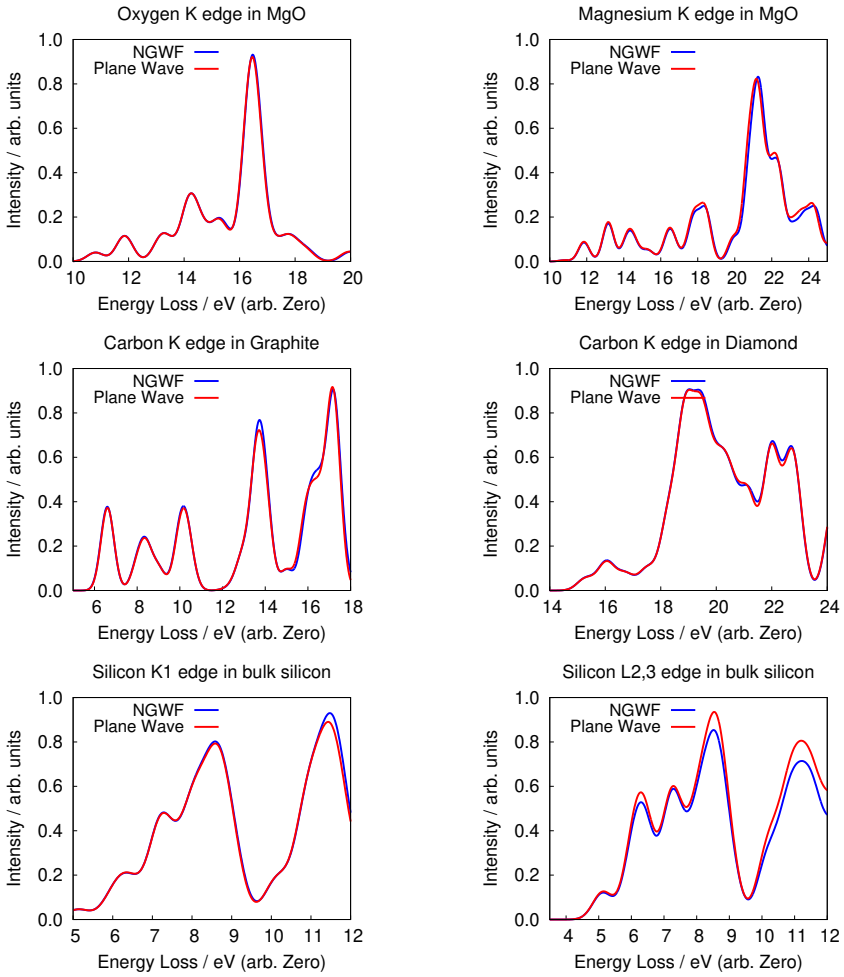
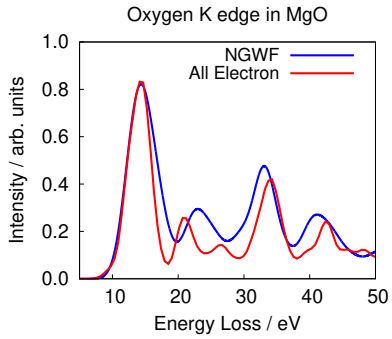
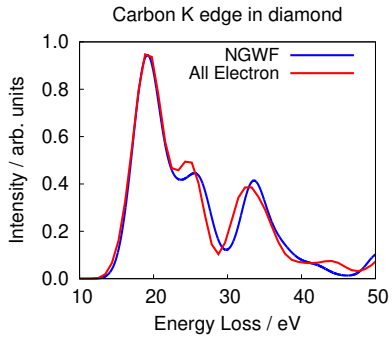


Figure 3.9: The predicted EEL spectra for the materials shown under both a conventional plane wave-PAW approach and using the ONETEP implementation. The spectra have been aligned by hand.



(a) MgO O K edge



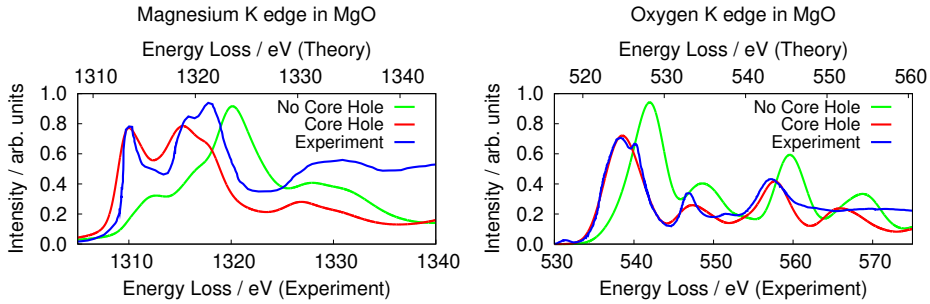
(b) Diamond C K edge

Figure 3.10: Comparison of results between ONETEP and ELK codes.

that the dipole approximation is also valid.

3.10.4 Comparison To Experiment

Previous comparisons to theoretical methods demonstrated the correctness of the new implementation within the bounds of simulation. Experiment provides a far superior benchmark. The experimental data have been rescaled to match heights with the first major peak in the predicted spectrum. A fixed broadening of 1.5 eV was used to compute the predicted spectra, as this value roughly matches the instrumental resolution in the experiments used for comparison. Through the experimental comparisons this broadening



(a) (b)

Figure 3.11: The magnesium and oxygen K edges in MgO. Here, the upper x axis measures energy loss as computed using the Mizoguchi method. The lower axis gives the experimental energy loss. For the Mg K edge, a different broadening scheme was used. Experimental data from [88].

scheme has been kept mostly consistent. For the MgO Mg K edge and the graphite C K edge however, a lifetime broadening scheme was used instead of the ordinary Gaussian smearing. This reflects the longer lifetimes (and thus narrower line widths) of the low energy peaks which are clearly visible in the experimental spectra. Zero momentum transfer was assumed. A full core hole in the 1s orbital was used.

From Figure 3.11, it can be seen that very close agreement with experiment can be obtained. A degree of work is needed to get such good agreement, mainly in the selection of the correct broadening scheme. For the materials studied, it is clear that a full core hole is sufficient to get a good match with experiment and thus one additional matching parameter (the fraction of a core hole to use) can be eliminated.

MgO Mg K

Here, we can see very clearly the impact of including a core hole in a simulation. Without the core hole, the predicted spectrum is dominated by a

peak at approximately 1325eV. This peak is shifted to lower energy and a second peak, which in the non core hole spectrum is little more than a shoulder, is raised in intensity. The combination of these two effects produces far better agreement with experimental results. The ratio of peak intensities in our predicted spectrum is explained to a degree by the choice of broadening scheme. An energy dependent Lorentzian correctly produces the sharp peak near the onset, but also reduces the intensity of higher energy peaks. This leads to a reversal in the heights of the peaks at 1312eV and 1320eV relative to experiment.

MgO O K

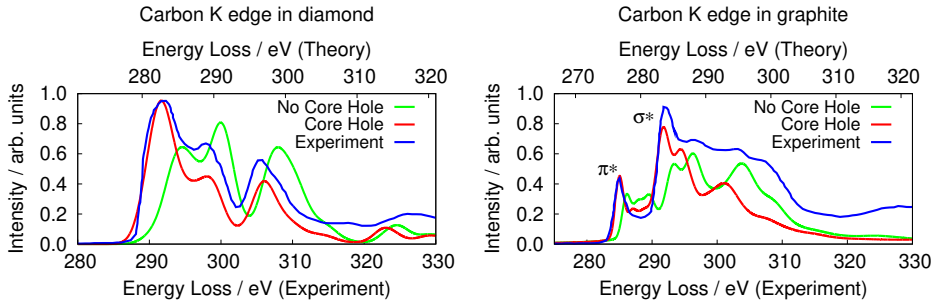
Here again, the core hole shifts the onset of the peaks to lower energy and reduces the intensity of the peaks somewhat. The small feature on the right of the first peak in the experimental spectrum is not reproduced. It is unknown if this is a result of an impurity, sensor defect or genuine feature of the material.

Diamond C K

Diamond is a wide band gap insulator like MgO and as such we expect a core hole to be important here. This reasoning is correct and inclusion of a core hole in the simulation shifts the spectrum onset to lower energies by about 3eV and also reverses the intensity pattern of the first two peaks, brining the observed ordering of intensities into agreement with the experimental spectrum. This is shown in Figure 3.12a.

Graphite C K

Unlike the previous material, graphite is a semi metal and much better screening of the core hole is expected. As a result, the non core hole spectrum already matches reasonably well with the experimental result. The degree



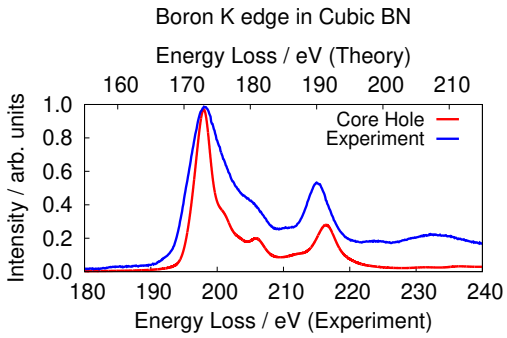
(a) (b)

Figure 3.12: The carbon K edges in diamond and graphite. Again, the upper x axis measures energy loss as computed using the Mizoguchi method. The lower axis gives the experimental energy loss. For the graphite C K edge, a different broadening scheme was used. Note the reproduction of the π^* and σ^* peaks. Due to the greater conductivity of graphite, there is better screening of the core hole and thus the explicit inclusion of a core hole in the simulation does not improve agreement with experiment as dramatically as for the wide gap insulator diamond. Experimental data from [118].

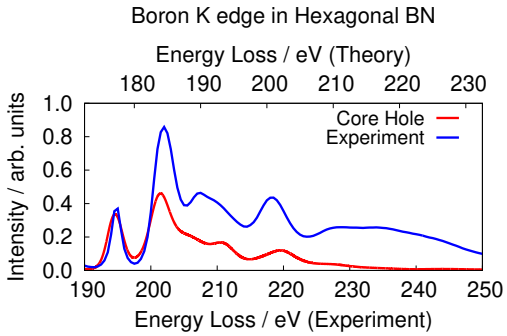
of agreement is still improved by the inclusion of a core hole. Graphite is also well known for its delocalised π system, which leads to a π^* peak at low energy losses. The peak intensity suppression for higher energies described in Figure 3.10.4 may also explain the discrepancy in the ratio of intensities between the π^* and σ^* peaks in the predicted spectrum compared to experiment.

3.10.5 BN

As good experimental spectra were available from the European EELS archive, an additional study was made of various BN poly-morphs. These were only compared to experiment and simulations included a full core hole.



(a)



(b)

Figure 3.13: The Boron K edges in cubic and hexagonal boron nitride, as well as in a BN nano tube. Experimental data from [117].

Cubic BN

Once again, good agreement with experiment is seen. The main feature of the experimental spectrum, the two peaks at 200 and 217 eV (experimental loss), are present in the predicted spectrum. The transition leading to the shoulder on the left hand side of the first experimental peak appears to be correctly captured by the simulation, as shown in Figure 3.13a. though at slightly too high an energy and intensity.

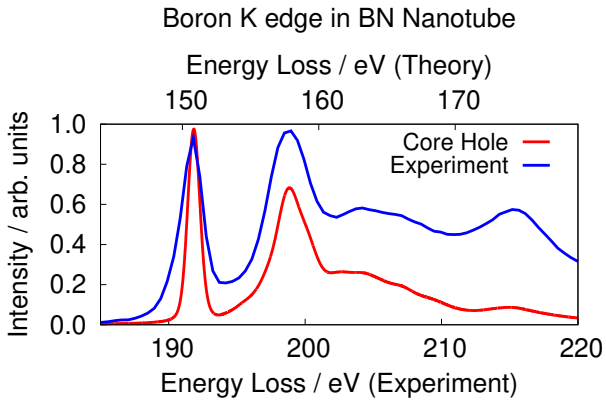


Figure 3.14: Comparison of simulated and experimental spectra for a boron nitride nanotube. In spite of the problems converging the conduction manifold, a good qualitative agreement is still achieved.

Hexagonal BN

There is a correspondence between diamond and cubic BN, and graphite and hexagonal BN. As expected from this correspondence, a strong π^* peak is seen at low energy and a σ^* peak is also present. These are seen in both experimental and predicted spectra. Subsequent peaks are also reproduced by simulation, though with reduced accuracy in intensity and relative energy.

BN Nanotube

This system provides the first test of the method on a nano-material, the sort of target system for which ONETEP is particularly suited. The BN nanotube also provides a particularly tough test of ONETEP's conduction optimisation feature, as the conduction manifold of the tube contains many tightly spaced states with little space in between. Careful choice of the number of states to optimise is needed in order to achieve convergence of the conduction NGWF CG minimisation. EELS simulation in nano-materials is not yet a 'black box'. Possible causes of this poor convergence include the presence of highly

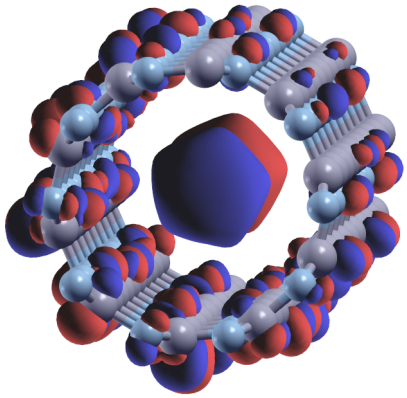


Figure 3.15: A conduction manifold state of the BN nanotube. The isosurface is drawn at the 0.01 level. This state is the 101st above the highest occupied orbital: 6.55 eV above the highest occupied orbital.

delocalised states which run down the centre of the tube. Figure 3.15 shows an example of such a state. As ONETEP uses an atom centred basis, the maximum amplitude for these states is spatially separated from the centres of NGWFs. This means that the optimisation routine must attempt to optimise conduction NGWFs to represent these delocalised states using only the margins of atom centred NGWFs. At first this might seem like a severe problem, as these states will never be well represented. However, it is key to remember that EELS only probes the conduction band wavefunctions in a region localised to the extent of an atomic core wavefunction. As a result, poor representation of this class of conduction band wavefunction is immaterial to the quality of the predicted spectrum. This may explain the good degree of agreement seen in Figure 3.14 despite the failure of the conduction optimisation.

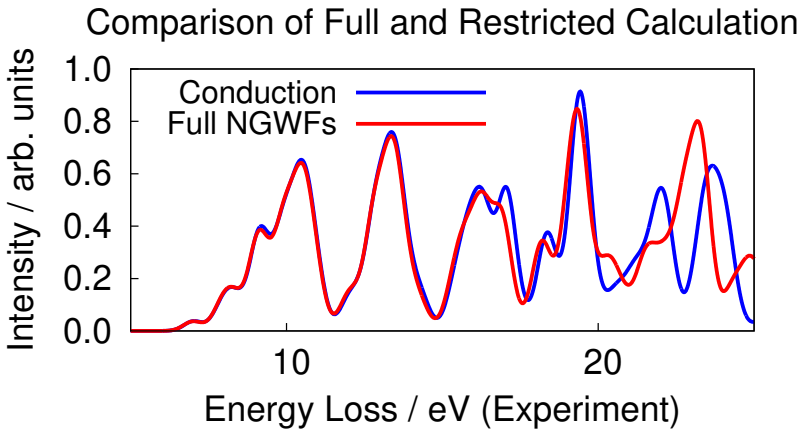


Figure 3.16

3.11 Reducing Effort for Conduction Optimisation

A series of tests were conducted to investigate if conduction optimisation and subsequent diagonalisation could be made more efficient by only including conduction NGWFs on atoms within a certain radius of the atom for which an EEL spectrum is being calculated. This idea builds on the problem discussed above, with the optimisation of conduction states for BN nanotubes. If EELS is only sensitive to the wavefunction in the vicinity of the atom being excited, then it is only necessary to represent the conduction manifold well in the region close to the excited atom.

Figure 3.16 shows the result of using this method in bulk silicon, where conduction NGWFs were included on atoms within 6.0Å of the excited atom. Comparison to the spectrum computed using conduction NGWFs on every atom shows that for the first fifteen eV of the edge, there is only a minimal difference between the two methods.

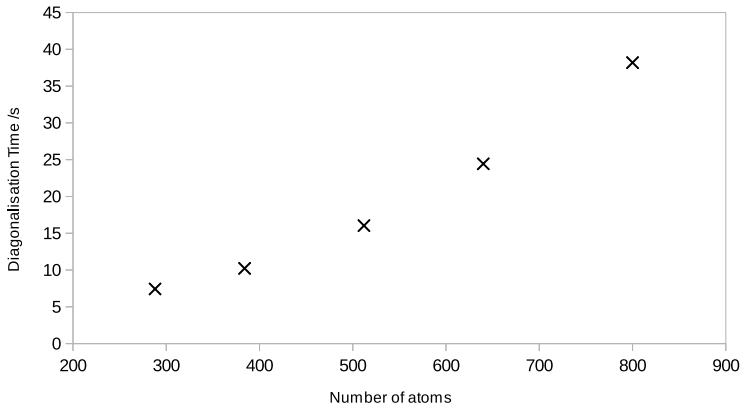


Figure 3.17: The time taken to perform the joint basis diagonalisation for diamond systems of increasing size. The superlinear scaling is expected.

3.11.1 Scaling Tests

The objective of implementing EELS simulation in ONETEP was to enable the prediction of spectra for systems which were previously too large for conventional DFT codes. To assess if this objective has been met, some simple scaling tests were conducted. As described in Section 2.2.1, an $\mathcal{O}(N^3)$ diagonalisation step is required to in order to perform an EEL calculation. Figure 3.17 shows that this is not a great barrier, even when dealing with the largest systems. The crossover point where diagonalisation would dominate the cost of a calculation is $\mathcal{O}(10000)$ atoms.

3.12 Conclusions

A means of simulating EEL spectra has been added to the ONETEP code. Tests against existing DFT codes, which use the same method for spectral simulation, confirm the correctness of our implementation. Comparisons to CASTEP show that differing choices of pseudopotential have minimal impact on well converged spectra. The pseudopotential and dipole approximations

were tested against an all-electron DFT code. Once again, good agreement was seen in this situation demonstrating that these approximations are robust.

Our method has been successfully validated against experiment. Inclusion of a core hole led to an improvement in the match between prediction and experiment in every case studied. In some cases, this improvement was dramatic. A full *ab initio* spectral prediction is still not possible, as empirical broadening parameters are required. A future project to support the adaptive broadening schemes available in OptaDoS would greatly improve the situation and remove the need for *post hoc* fitting.

Throughout this investigation, difficulty was occasionally encountered in converging conduction NGWF calculations. It certainly appears that calculations are not as robust as ground state calculations, which converged quickly (fewer than 20 NGWF conjugate gradient steps) in every case. Often, the causes of this difficulty could be traced to particular oddities of the system and may be anticipated, for example in situations where conduction states might have significant amplitude in areas distant from atoms. In other cases, inspection of the (inaccurate) valence calculation derived conduction band DoS may show a continuum of states, which resist the LNV based approach to conduction optimisation.

Going forward, this methodology promises to be useful in the investigation of large and complex nano-structures. To demonstrate this the newly implemented EELS functionality will be used to investigate a defective anatase surface in the following chapter.

Chapter 4

An Application of LSDFT EELS To Oxygen Vacancies In Anatase

4.1 Defects in Anatase

Anatase has a number of promising features as outlined in the introduction to this work. Many applications exploiting these features are heavily impacted by the presence of defects which can, for example, act as charge recombination centres, or sites for catalytic activity, or charge trapping. When studying a material, we are thus very interested in identifying which defects are present and assessing their concentration. Once the prevalent defect species in a material is known, it is possible to use both experimental and theoretical methods to assess what impact it might have on performance in a given application. Electron energy loss spectroscopy is a very powerful technique, which is especially sensitive to local coordination and geometry. These characteristics are expected to be altered near a defect. As a result, EELS is a promising method for studying defects. This is especially true when the high spatial resolution of EELS is considered, which allows point

defects to be studied.

Ultimately, the objective of this study was to test if localised defects in anatase could be identified with EELS. The variation of EEL spectral structure with distance from defect was investigated along with how a finite sample thickness would change the spectrum by contributing a background signal.

The oxygen 2+ vacancy was selected for study as it is one of the better studied defects in anatase [119, 120]. The 101 facet was chosen as it has the lowest surface energy of the anatase facets and thus makes up the majority of the surface of an anatase nanoparticle.

4.2 Prior Work

EELS has been used to study a range of insulating systems, including oxides. Studies have also been conducted on strontium titantate in the hope of being able to extract the contribution of the surface to the overall EELS signature [121].

Early work on titanium oxide materials demonstrated the ability of EELS to distinguish between the anatase and rutile phases of TiO_2 [122]. More complex phases have also been studied, with a view to extracting information about the coordination environment present in the material [123]. Multivariate analysis methods have been applied to studying the spectra of reduced TiO_2 surfaces [124]. More recently, Zhang *et al.* have used EELS to study the interface between anatase and a nanotube within a complex core shell nano structure [125]. Their results are interesting as the changes in carbon K and titanium $L_{2,3}$ edges at the interface are observed, indicating that EELS may be used to study interfaces in nano structures. In situations like this, the ability to simulate EEL spectra may be particularly appealing as this would provide a further means of establishing whether a proposed interface structure is correct.

Other oxide materials have also been studied. These include gallium oxide, where EELS was used as a sensitive probe of the local coordination environment [126].

Experimentally, EELS has been used to study technologically important processes such as the insertion of lithium into iron phosphate matrices used as battery electrodes [127], barium titanates [128] and ϵ-Fe_2O_3 [129].

Single atom substitutional defects in graphene have also been studied [130], with graphene systems having the advantage that the background signal from other atoms can be eliminated when the sample is analysed using the electron beam normal to the sample surface. Whilst measuring dopant concentrations and segregation via elemental mapping is widespread, the use of EELS to investigate vacancies is less common.

4.3 Model Systems

4.3.1 The (101) Surface

A model of the (101) anatase surface was constructed using DFT lattice parameters. An annotated image of the slab (without the vacuum gap) is shown in Figure 4.1. The model system consists of 720 atoms in a double sided slab and a 36Å vacuum gap. This size system was selected both to exhibit the capacity of ONETEP for tackling large systems and to provide adequate space to accommodate any long ranged relaxations around defects. Furthermore, a large surface slab provides a good choice of sites to perform EELS simulation and allows the probing of atoms a long way from the defect. As an initial step, the model system was relaxed using the ONETEP implementation of the BFGS algorithm [131] until residual forces were less than 0.1 eVÅ^{-1}. After this, a surface bridging oxygen atom was removed and the structure subjected to a further round of geometry optimisation. During this second round, a tight condition on energy convergence for each NGWF SCF step

was imposed in order to ensure that well converged forces were obtained. Structures were deemed to be relaxed when the residual energy per atom, forces and maximum displacement were less than 1×10^{-6} Ha, 1.1×10^{-3} Ha Bohr^{-1} and 5×10^{-3} Bohr respectively.

The introduction of two surfaces to the model system provides the opportunity for structural relaxation in the direction normal to the slab. From this it would be expected that the slab system should show elongation of bonds normal to the surface as the whole slab relaxes to a thinker configuring. This is not seen, possibly because expansion is still constrained in the other two directions and no bonds are exactly normal to the surface - meaning any relaxation in the direction normal to the surface would attract an energy penalty resulting from compression in the other two directions. Quantitatively, for comparison, DFT bond lengths in bulk anatase are between 1.95Å and 2.00Å, while in the relaxed slab they vary from 1.78Å to 2.11Å.

For the computation of EEL spectra, a full core hole in the oxygen 1s orbital was used. Anatase is an insulator and so core hole effects are expected to be important. Once again, the PBE functional was used with an 800eV cut-off energy and $10a_0$ NGWFs. The JTH [108] PAW dataset was used. For the production of spectra from matrix elements, OptaDoS was used with a Gaussian broadening of 0.6 eV combined with a Lorentzian width of 0.2 eV and a Lorentzian scale of 0.1.

For all calculations the PBE functional[37] was used. This functional was selected as it provides an acceptable trade-off between speed and accuracy. In transition metal oxides DFT+U methods are frequently used. These methods introduce an additional term to the DFT Hamiltonian accounting for the correlation of d electrons often seen in these materials. In the present study this technique was not used as it was felt it would introduce an additional tunable parameter (the value of the Hubbard U term) to the spectral prediction process. There is however good reason for considering the introduction of a U term in future studies as the lower energy states of the anatase

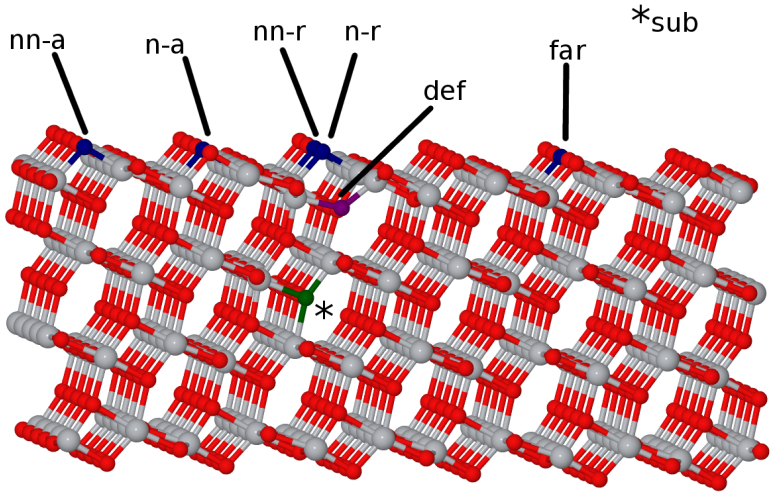


Figure 4.1: The anatase surface cell used in the study. Sites of atoms for which EEL spectra were calculated are indicated. The atom shown in purple and marked `def` is the closest oxygen atom to the defect centre.

conduction band have a significant Ti-d character.

Notation

Defects are indexed by their depth into the cell using the convention shown in Figure 4.1. Atoms in the same row (along the direction pointing into the page in figure 4.1) as the missing oxygen atom in Figure 4.1 are labelled `n-t` and `nn-r` the neighbouring and next nearest neighbouring atoms respectively. Atoms in the same row (left-right) as the missing atom are labelled `n-a` and `nn-a` for the neighbouring adjacent and next nearest neighbouring adjacent atoms respectively.

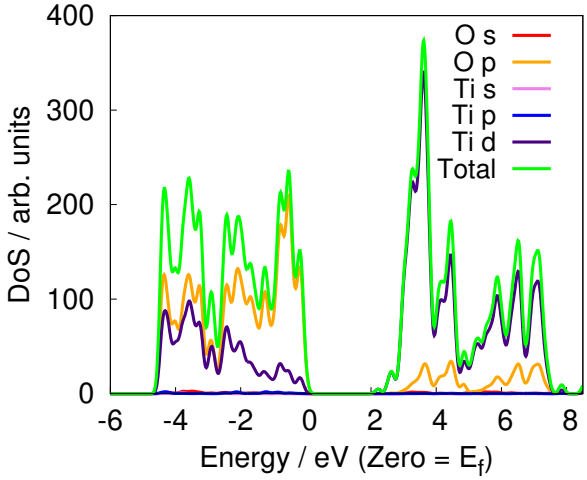
Defects are notated using the convention $V_{species}^{charge}$ for a vacancy. This study looks at V_O^{2+}. Later chapters will also examine V_O^{1+} and V_O^0.

4.4 Results

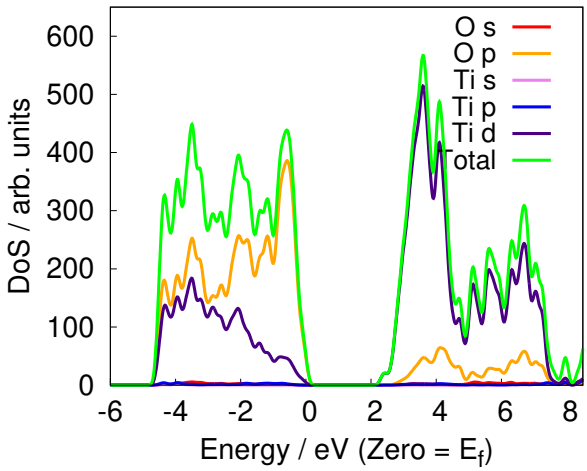
4.4.1 Electronic Structure

Figure 4.2 shows a comparison in density of state between bulk and surface systems. In both cases the density of state has been decomposed by both species and angular momentum channel. For both oxygen and titanium there is a negligible contribution from the s (zero angular momentum / spherical symmetry) channel. The valence manifold if composed largely of oxygen p orbitals, with some contribution from titanium d states. The conduction band, especially at the band edge, is composed almost entirely of titanium d orbitals, with a small contribution from oxygen p orbitals a higher energy. These features are common to both bulk and slab systems. There are, however, some differences between the two. The most notable difference is the change in the density of states at the bottom of the conduction band. Both systems have a 'double peak' between 2-5 eV. In the slab system the intensity of these peaks is roughly equal, whereas in the bulk system the first has around twice the amplitude of the second. There is also a more subtle change in the valence band, with a larger density of states at the upper edge of the band (relative to the rest).

To determine the degree to which the electronic structure of the slab system approaches that of the bulk in the middle of the slab, projected densities of states for each layer of the slab were computed. Here a 'layer' is taken to refer to a plane of titanium atoms with (approximately) the same z coordinate (z direction is normal to the slab surface) and a stoichiometric number of oxygen atoms bonded to those titanium atoms. In all the model system used contained four layers, two surface layers on the top and bottom and two layers in the middle. By symmetry about the centre of the slab these four layers are reduced to two distinct environments - 'top' representing the exposed surfaces and 'middle' representing the interior of the slab. The contributions of these two environments is shown in figure 4.3, with the bulk



(a) Bulk anatase pdos



(b) Anatase slab pdos

Figure 4.2: The density of states for anatase systems, the bulk is shown in the first figure and the surface in the second. Contributions to the density of states have been divided by species and further divided by angular momentum, which is a proxy for the symmetry of the orbitals making up the states.

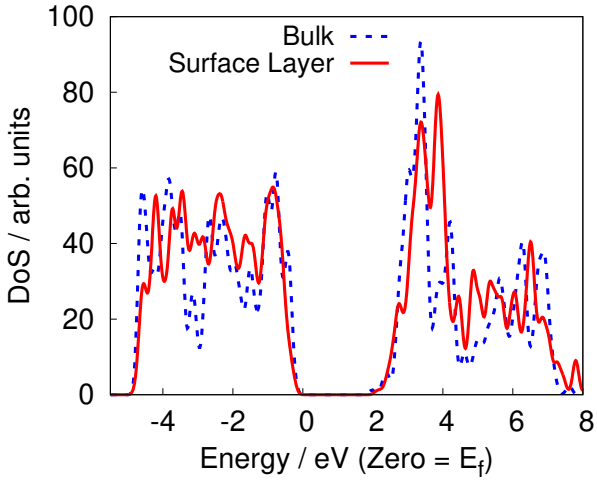
density of states underlaid as a dashed line for comparison.

Figure 4.3a shows notable differences between the bulk PDoS and the PDoS for the 'top' or surface environment - the strong double peak at the lower edge of the conduction band is apparent in the latter. Differences in the valence band are less apparent. The PDoS for the interior of the slab reveals fewer differences with the bulk PDoS, the only notable feature being that there is a less pronounced peak at the lower edge of the conduction band than is present in the bulk. The double peak structure seen in the surface PDoS is reduced to a single peak with a shoulder - showing that the even one layer into the slab the electronic structure is already returning to bulk-like.

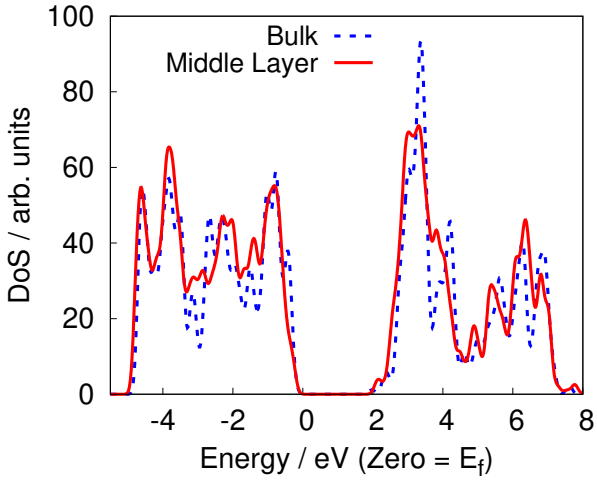
4.4.2 Comparison to experiment

Figure 4.4 compares the EELS spectrum of anatase powder with that obtained from [132]. Due to the low experimental resolution (0.7 eV) agreement of fine structure is impossible to ascertain but the general morphology of the experimental spectrum is correctly reproduced with the exception of the second double peak. A hint of a shoulder can be seen in the experimental spectrum, indicating that the intensity is incorrect but the energy of the transition relative to the onset is correct. The edge onset of the predicted spectrum was calculated using the Mizoguchi method.

Figure 4.5 shows an EEL spectrum for the perfect (101) surface, compared to an experimental X-ray absorption near edge structure (XANES) spectrum for the surface. Whilst EELS and XANES are two distinct methods, both probe the same transitions and thus involve the same matrix elements. As a result, the comparison between the two provides insight into the accuracy with which we have modelled the conduction band states in the vicinity of the surface. A good agreement is obtained, providing some confidence in our model.



(a) Anatase slab surface layer pdos vs bulk



(b) Anatase slab middle layer pdos vs bulk

Figure 4.3: Density of state for anatase slab resolved into layers. The first figure shows the contribution from the top surface, the second shows the contribution from a middle layer. In both cases the DoS for the bulk system is included for comparison.

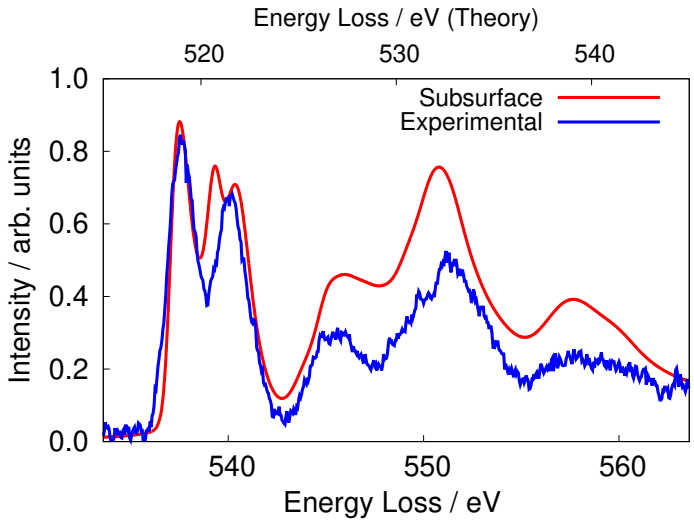


Figure 4.4: A comparison of an experimental EEL spectrum from [132] compared to a simulated spectrum for the subsurface (`sub`) oxygen atom in the surface.

4.4.3 Slab Vs Bulk

An EEL spectrum for bulk anatase can also be computed. To produce figure 4.6 a simulation cell measuring 10.43Å×15.19Å×24.77Å containing 336 atoms. The DFT lattice constants were used to create the simulation cell. A whole core hole is the oxygen 1s orbital was used. Figure 4.6 shows a good degree of similarity between simulated and experimental results. The major discrepancy is in the intensity of the two sub peaks of the double peek between 537 and 542 eV (experimental). The experimental spectrum shows a greater intensity for the first peak, while in the predicted spectrum the second peak is more intense. This may be caused by the choice of life-time broadening used to generate the predicted spectrum - a higher value would tend to suppress peaks at higher energy and could flip the relative intensities. The bulk EEL spectrum can also be compared to XANES data such as that of Mosquera *et al*[133]. As with the experimental EEL data a reversal of

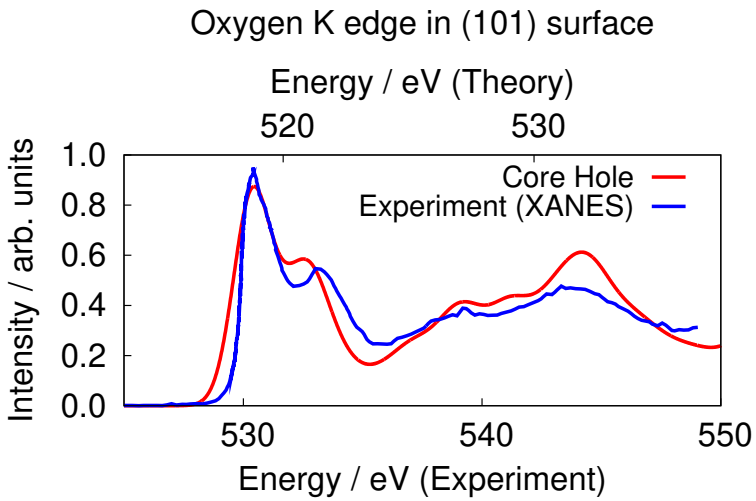


Figure 4.5: A comparison of a simulated EEL spectrum for an atom in the anatase surface with an experimental XANES spectrum [134]. While these two techniques differ greatly in their experimental realisation, they probe identical transitions and thus at a DFT level of theory are modelled identically.

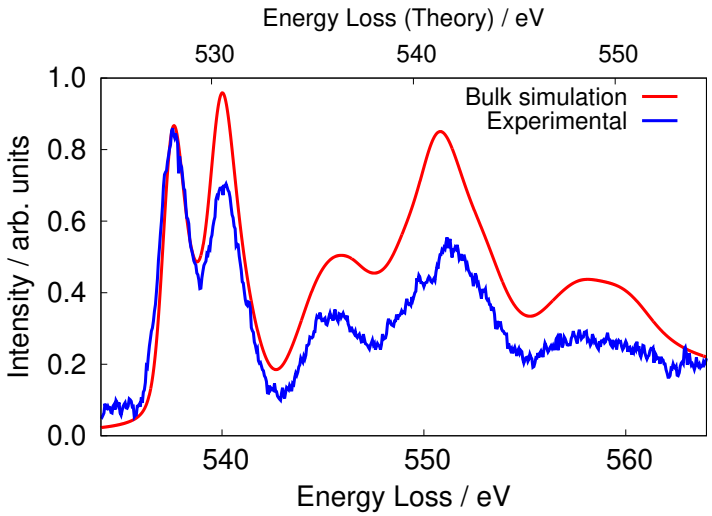


Figure 4.6: A comparison of an experimental EEL spectrum from [132] with a simulated spectrum for an oxygen atom in bulk anatase.

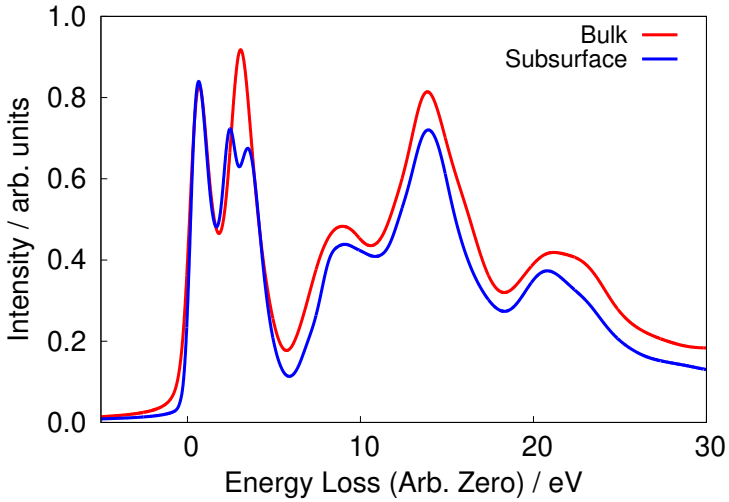


Figure 4.7: A comparison of a simulated EEL spectrum for a subsurface oxygen atom with a simulated spectrum for an oxygen atom in bulk anatase.

intensity in the first two peaks is seen.

Figure 4.7 provides a comparison between the predicated EEL spectrum for the subsurface oxygen atom in the slab system and an oxygen atom in bulk anatase. The prime difference between the two spectra is the breaking of the second peak into two sub peaks in the slab system. This may be attributable to the symmetry breaking effect of the surface in the slab system. Beyond this difference the two spectra show remarkable similarity, with only minor variations in peak heights and trough depths between the two. If the peak splitting seen in the subsurface atom arises form symmetry breaking then the phenomenon may persist for some distance into the slab.

4.4.4 Surface Vs Interior

As an initial test oxygen K edge spectra were computed for two atoms in the perfect anatase surface model. One atom was a surface bridging oxygen. The second was in the interior of the system, but had a similar coordination as the

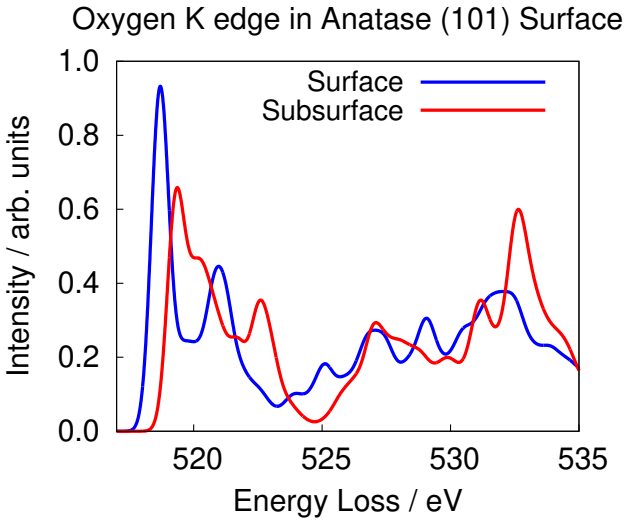


Figure 4.8: Comparison of the spectra for an oxygen atom in the perfect surface and one in the subsurface region.

surface atom. The results are shown in Figure 4.8. There are differences in the two spectra. Notably in the subsurface atom, the energy onset has shifted approximately 1eV higher and the first peak has broadened and reduced in intensity. A shoulder on the first peak is also apparent. At higher energies there is a greater difference in the spectra, but this is unlikely to be accessible to experiment as lifetime broadening effects will obscure these peaks.

The double peak seen in experimental EEL spectra of anatase is also present, though with an intensity ratio which differs significantly from that seen in experiment. Based on experience in Chapter 3, this may be due to core hole effects. The band gap in anatase is not as wide as in diamond and MgO, and so some level of screening will take place. This may mean that some fraction of a core hole would better represent reality. There is a risk of over fitting in situations like this, and the decision to use a full core hole was retained.

If a series of spectra for different sample thickness are available it may be

possible to resolve contributions to the spectrum from surface and subsurface atoms. To assess if this is the case an independent component analysis (ICA) based method was used to attempt resolve the contribution of the surface to a spectrum. This idea is inspired by the work of Zhu and Botton [121] who used a least squares fitting method to extract a surface contribution from a thickness series. ICA has also been applied in the field of EELS analysis previously, for example to map elements with overlapping edge onsets [135].

To test this idea, a series of spectra were produced using the combination method discussed in Section 4.4.6 with differing surface and bulk contributions. Gaussian noise was also included. The Scikit-learn package was then used to perform an ICA. The decomposition and input spectra are shown in Figure 4.9. The greatest surface contribution was 15%, corresponding to a sample consisting of 5 layers. The bulk contribution to the spectrum was easily recovered. The second component bears a strong resemblance to a negated signal from the surface. This indicates that some meaningful information about the surface spectrum of anatase may be recoverable from sufficiently low noise spectra of samples with differing depths.

An attempt was also made to determine if ICA could recover the defect signal against a background of bulk like oxygen K edge signals. Unfortunately, in this case, even assuming access to unrealistically thin depth series it was not possible to extract a signal corresponding to the defect spectrum.

4.4.5 Distance from defect

In the defective surface model, a selection of oxygen atoms at different distances from the defect site were selected (Table 4.1) and their EEL spectra computed. For later work combining spectra, the Mizoguchi energy offset was computed for each spectrum. The resulting edge offset energies are also supplied in the table. Note that while these differences may be experimentally resolvable, they also carry uncertainties of the same magnitude as the difference between them owing to system size convergence effects discussed

in Section 3.9.

Atom	Distance to Defect /Å
`def`	0.0
`n-r`	4.46
`nn-r`	8.01
`n-a`	6.36
`nn-a`	10.96
`far`	12.54
`sub`	4.00

Table 4.1: Distances to the relaxed defect atom of the oxygen atoms whose K edges were computed.

Figure 4.10 shows the spectra computed for the atoms near the defect. The nearest atom, `def`, is the most distinctive of the near surface atoms, showing a shift in edge onset towards higher energies of 0.5 eV and a reduction in the shoulder structure on the right hand side of the peak when compared to spectra to more distant atoms. The large change seen for this atom should not be a great surprise. Examining Figure 4.1 shows that the `def` atom relaxes by a significant amount, lifting up from its previous position towards the surface. This changes its local coordination environment; specifically it is now closely coordinated by two titanium atoms instead of three. In this respect it becomes more surface like, however the coordination is closer to linear than the surface bridging structure. The sensitivity of EELS to coordination thus explains the difference.

The other surface atoms show very little change in their spectra, especially at low energy losses. The spectrum of the `far` atom differs little from that of a surface atom in the defective slab. This confirms that EELS is a short range method in this material. This result may again be explained by the minimal change in coordination of these atoms.

The spectrum for the subsurface atom shows a different pattern of peaks. The first peak is split into two, with the second of those showing further subdivision. At energies of approximately 532 eV, there is an increase in peak intensity in the sub surface spectrum relative to the near defect surface atoms.

4.4.6 Spectra From Multiple Atoms

To account for the finite width of the probe electron beam and to model the excitation of multiple atoms through the thickness of the sample, a simple scheme was adopted. A weighted average of two or more spectra was taken. This process makes the assumption that EEL transitions of separate atoms may be treated in isolation. As the lifetime of such transitions is very short, it is likely that this approximation is a good one.

The mixing process is simply done by taking a weighted combination of the two spectra:

$$S_{\text{combined}} = \alpha S_1 + (1-\alpha)S_2\,, \quad 0 \leq \alpha \leq 1\,. \tag{4.1}$$

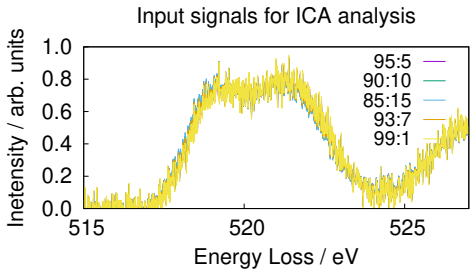
Figure 4.11 shows the results of applying this process to in two locations on the slab. The first is directly on top of the defect where the spectra for the `def` and `sub` atoms have been mixed. In the second, the `far` and `sub` atoms have been mixed. Examining the figure indicates that the split first peak encodes information about the presence of a defect. The spectrum for a defective column tends to a flat top instead of a double peak and a valley only develops between the two peaks when a large contribution from the bulk is present. In columns of atoms lacking a defect, a valley between the first peaks is always present and tends to diminish (rather than grow) as sample thickness is increased.

4.5 Conclusions

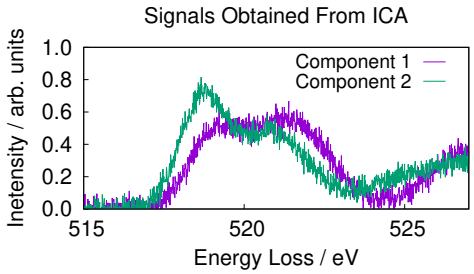
Atoms near oxygen vacancies produce distinct EEL spectra. The use of these spectra for identifying defects is likely to be inhibited by sample thickness effects, whereby the contribution of a small group of atoms near a defect is likely to be overwhelmed by atoms in a more bulk-like environment which are also excited by the beam. It is thus likely that detecting vacancy defects in anatase surfaces using EELS will prove challenging, though not impossible, with current technology. It is conjectured that the most promising systems for study are grain boundaries, where whole columns of atoms with the same (though different from the bulk) coordination may be probed. Native interstitial defects, where the immediate coordination environment of the interstitial atom is likely to differ greatly from the bulk environment and thus lead to a significantly altered EEL spectrum may also prove easier to identify against a background of bulk signals.

Non native doping, of any kind, may be readily identified as the edges associated with the dopant atom will not be present in the bulk material. In this case, EEL modelling can still prove useful, as a number of possible coordination geometries can be simulated and the results matched to experiment.

An avenue for further work would be the extension of this method to support spin-orbit coupling, which is in the process of being introduced into the ONETEP code. This is of particular relevance to anatase surfaces as the Ti L2,3 edge is relatively easy to access experimentally and is known to be a sensitive indicator of the TiO_2 phase.



(a) The input signals, with noise included. There is little to tell these apart.



(b) Using ICA, the signals indicative of the bulk and surface spectra can be recovered.

Figure 4.9: ICA can recover surface signals from combined spectra blurred with noise. Comparing the results here to Figure 4.11b we can see that the ICA extracts the key difference between the surface signal and the bulk signal, the former has a sharp peak at the onset, while the latter has a flatter double peak structure. Thus we identify component 1 with the bulk and component 2 with the surface.

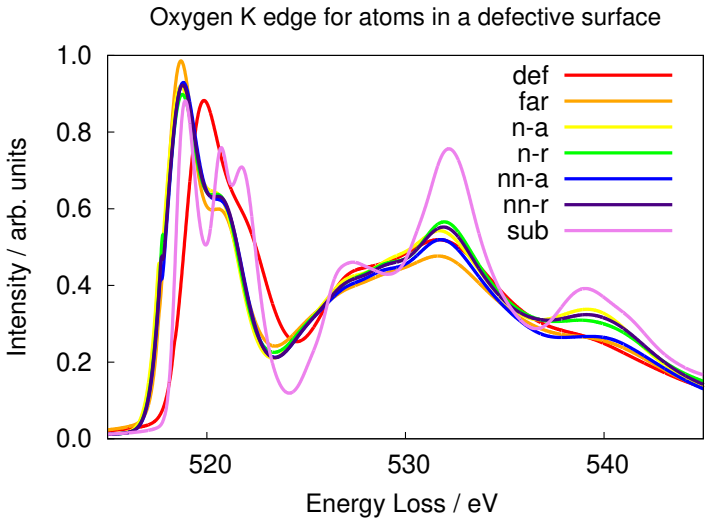
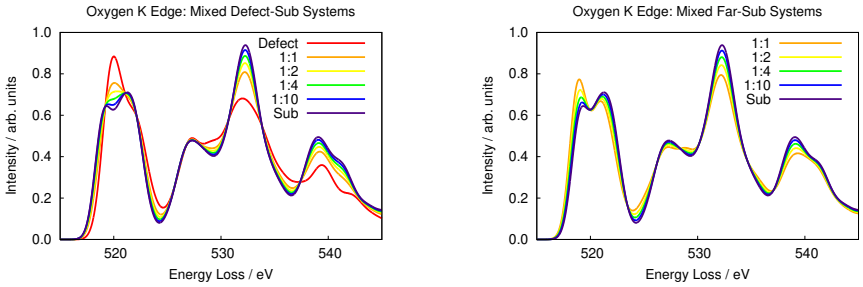


Figure 4.10: The spectra obtained from the indicated atoms in Figure 4.1. Most of the surface atoms are sufficiently distant from the defect to show no great change in spectrum.



(a) The spectra resulting from combining `def` and `sub` atom signals in different proportions. The labels are the `def`:`sub` ratio

(b) The spectra resulting from combining `far` and `sub` atom signals in different proportions. The labels are the `far`:`sub` ratio

Figure 4.11: Combined spectra to simulate an electron beam interacting with multiple atoms.

Chapter 5

An Application of LSDFT EELS To Nitrogen Doping in Anatase

Nitrogen is easy to incorporate into the anatase structure and many groups are investigating the resulting material, both experimentally and using simulations [5, 136–138]. Doping with nitrogen is observed to enhance the visible light photocatalytic activity of anatase [139]. Enhanced visible light activity is a very desirable property. As such photocatalysts would likely show higher efficiency when used under solar illumination. The generation of states in the band gap is associated with doped samples having a yellow tinge, instead of the pure white of intrinsic anatase.

Several different structures have been proposed for the inclusion of nitrogen in the anatase crystal structure. These include substitution for both oxygen and titanium as well as interstitials, with the first of these being the most common [140–142] . It would be useful to identify how nitrogen incorporates into the anatase lattice. Electron energy loss spectroscopy provides a means to identify how an individual dopant atom has incorporated. The local nature of STEM EELS would be useful in this case as dopant concen-

trations are around 0.25% on an atom basis [5] and achieving a high signal to noise ratio will require the electron beam directly striking the defect atom.

To achieve this, we simulate several different proposed defect structures in anatase [140] and compute EEL spectra for them. Using these spectra we show that it is possible to distinguish between different substitutional defects using EELS. This work represents a realisation of the original objective of implementing EELS in ONETEP: the generation of theoretical insight of utility to experimentalists. The nitrogen doped anatase system is ideal, with the dopant having a strong edge in a region free from the edges of other elements.

In their work, Chen and Dawson examine three possible geometries for the incorporation of nitrogen into anatase: nitrogen substituting on an oxygen site (N_O), nitrogen substituting on a titanium site (N_{Ti}) and a nitrogen interstitial (labeled $(NO)_O$ by Chen and Dawson, but could also be labeled as N_i. We adopt the former convention). The geometries and computed densities of state are shown in Figure 5.1. Chen and Dawson find that nitrogen substituting on a titanium site is the most energetically favorable. This result may be counter-intuitive, as the closer electronegativities of oxygen and nitrogen might suggest that the most stable site for nitrogen incorporation would be an oxygen site. Examination of the geometries obtained by Chen and Dawson show that when incorporated at a titanium site, nitrogen tends to produce a bound nitrate (NO_3^-) structure which, given the stability of the nitrate anion, may account for why the N_{Ti} substitutional is energetically favorable.

Here, the objective is to go beyond the more common use of EELS where the distribution of dopant atoms is mapped by looking for edges characteristic to the particular dopant atom. Instead, by examining the detailed structure of the edge, the local geometry of those dopant atoms may be obtained, giving further insight into how they have been incorporated into the host material.

5.1 Methods

The same computational parameters as used in the anatase surface study were used here. The model defect systems were constructed using the geometries of [140] as a guide. A bulk anatase super cell with a cuboidal shape using DFT-optimised lattice parameters was constructed. It contained 600 atoms and measured 18.98×18.98×19.42 Å. A nitrogen substitutional on an oxygen site, a nitrogen interstitial and a nitrogen substitutional on a titanium site were all prepared. These structures were subject to geometry relaxation. Final forces were less than 0.1 eV/Å.

Defects were studied in their -1 states, as [140] found these to be most stable over the upper half of the gap and these states showed no tendency to become spin polarised. All defects show large differences in their local atomic environments, which provides a good indicator that EEL spectroscopy may be able to distinguish between them.

The Nitrogen K edge (about 400 eV) was selected for this investigation as this is well separated from both the Oxygen K edge (around 530 eV) and the Ti $L_{2,3}$ edge (around 460 eV). This means that it should be possible to identify the spectral contribution from N atoms in a sample without interference from the bulk.

5.2 Results

Following ground state optimisation, the conduction band was analysed and a suitable range of conduction states was selected for optimisation. In all three cells this was approximately 2000 states (corresponding to 1000 bands) and corresponds to an energy range of about 7 eV above the top of the valence band. Initial inspection of the conduction manifold showed differences between all three defect types, which is consistent with the findings of [140].

5.2.1 Geometry

Figure 5.2 gives the final structures for the proposed defects. In the case of N_{Ti} in the negative charge state, we see the formation of a nitrite-like structure with a dangling oxygen atom, consistent with Chen and Dawson's results and those of experimental methods such as those of NMR, which detect oxidised nitrogen states after doped anatase has been prepared by calcination [143]. The distance to the dangling oxygen atom is 1.23Å. The nitrogen interstitial adopts a symmetric bridging structure, where it is bonded to two titanium atoms and also a central oxygen atom (Figure 5.2c). The nitrogen oxygen bond length in this case is 1.39Å. Finally, the N_O substitution relaxes into a structure which differs little from the non defective material, with the nitrogen not migrating from the original oxygen site and bond lengths to neighbouring titanium atoms changing by less than 3%.

5.2.2 Electronic Structure

Figure 5.3 depicts the density of states for the three defective systems. In the case of the nitrogen-oxygen substitutional defect there is minimal change in the electronic structure near the top of the valence band and bottom of the conduction band, with the most notable contribution being a set of states around 6 eV below the top of the valence band. It is perhaps not surprising that there is little change in the electronic structure as nitrogen is chemically similar to oxygen.

In both nitrogen-titanium substitutional and the interstitial additional states are seen at the top of the valence band, in addition to states at the bottom of the valence band. The additional states at the top of the valence band may be attributed to the more substantial electronic differences between nitrogen and titanium in the case of the substitutional and the profound structural changes associated with the interstitial.

Comparison between figure 5.3 and figure 5.1 reveals some significant

differences. These can, to an extent, be attributed to the different choices of functional: the present work used PBE, a GGA functional while Chen and Dawson used HSE06, a range separated hybrid.

5.2.3 EELS

The results for calculations without a core hole are shown in Figure 5.4. It is clear that all three defects produce very different EELS signatures. The N_{Ti} defect is characterised by a strong initial peak, while the other two defects show a lower initial intensity. The N_O defect shows two peaks at approximately 4 and 5 eV above the edge onset. The interstitial, however, produces a large signal at around 7 eV above the onset. This peak does overlap considerably with a peak from the N_{Ti}, which may interfere with its use in defect identification. Fortunately, the area closer the edge onset is still sufficient to distinguish the defects.

Having demonstrated, at least in principle, that it is possible to see significant changes in predicted spectrum, we then proceeded to produce predicted spectra which are directly comparable to those obtained from experimental results by including a core hole in the calculation. In this case, a full core hole in the nitrogen 1s orbital is used.

The method of Mizoguchi [106] was used to compute theoretical edge onset enrgies as discussed in section 3.7. Spectra for all three defects have been using the same x axis scale. Of imediate note is that the N_{Ti} defect shows an edge onset almost 5 eV higher than the other two defects. All three defects show a similar double peak structure close to the edge onset (around 395-400 eV in Figure 5.5). Even though there is some gross similarity in the spectra, the spectrum from the nitrogen interstitial ($(NO)_O$) shows a dip in intensity at around 397 eV, corresponding to a narrower initial peak. The N_{Ti} defect leads to a more pronounced shoulder on the second peak, at around 403 eV. The N_O defect is unique in that its two initial peaks show a 1:1 ratio, compared to the other two defects, which have an approximately 2:3 ratio in

the case of $(NO)_O$ and 1:2 for N_{Ti}. The most pronounced difference comes at around 407 eV, with the N_{Ti} defect showing a large peak in a region of the spectrum where the other two defects generate little signal. Examination of the projected density of states produced by Chen and Dawson in Fig. 5.1 reveals some nitrogen states in the conduction band. These may be repsonsible for this large intesnity peak.

5.3 Discussion

The great difference in the predicted EEL spectra for these three defects indicates that EELS analysis of a sample would be an effective means of determining which defect was present in the material. The high spatial and energy resolution available in modern instruments should allow a study of these defects in a nano material to determine where a nitrogen dopant sits in the crystal lattice.

5.4 Conclusion

The objective of this chapter was to establish if theoretical EELS could signpost the way to identifying defects in an oxide material by showing whether different defect structures could have sufficiently different EEL spectra that they could be distinguished using experimental data. It is clear from the above that this should be possible in the case of nitrogen doped anatase. It is hoped that this demonstration may inspire experimental investigators to search for the signatures of the defects outlined above. Furthermore, even if the GGA geometries and relative energies of the defects are not accurate, there is nothing to prevent the construction of other candidate geometries using, for example, hybrid functionals, DFT+U or higher order quantum chemistry methods. These geometries could then be used for EELS prediction in a larger supercell using ONETEP.

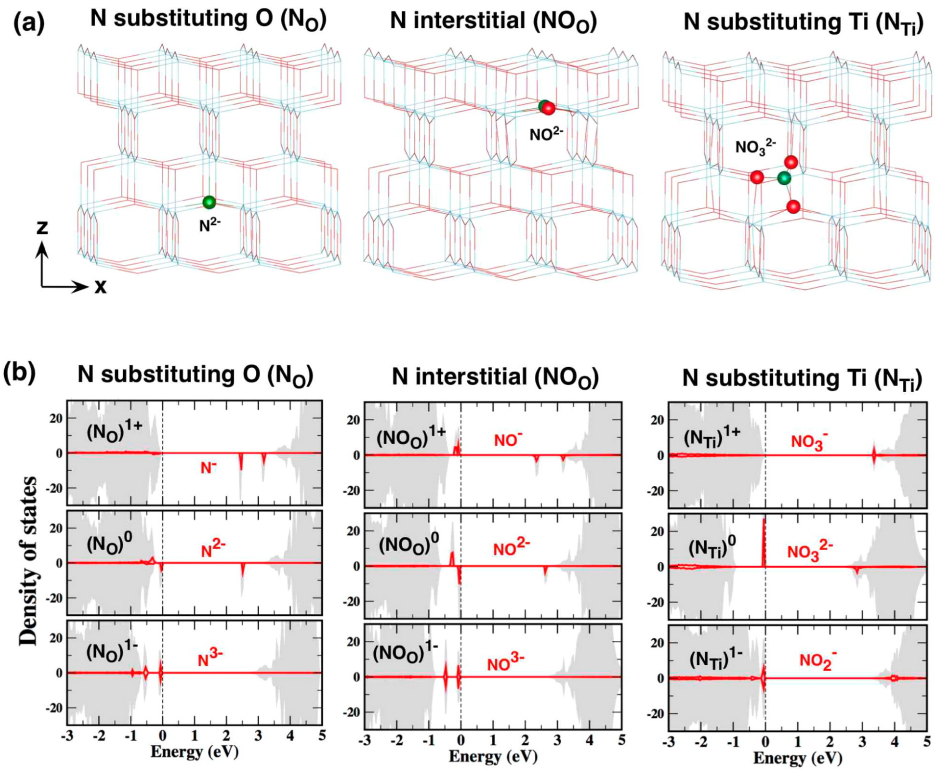
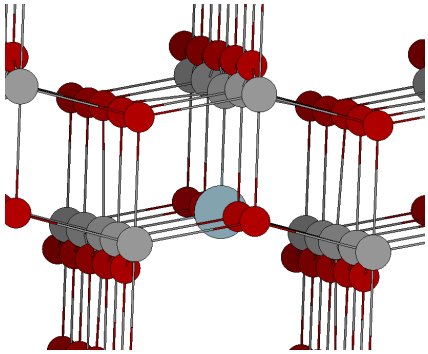
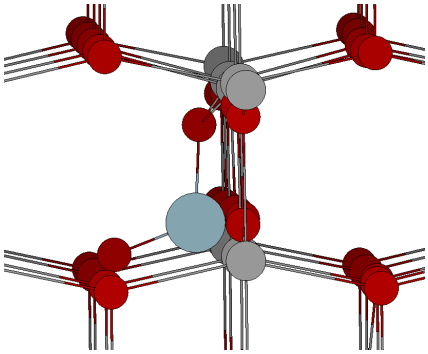


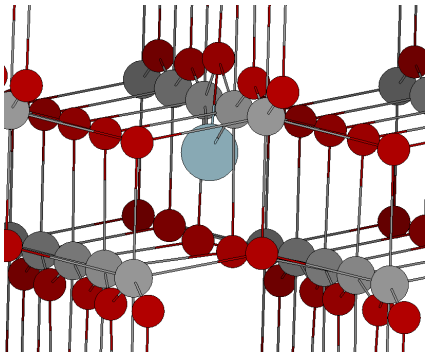
Figure 5.1: The three nitrogen incorporation configurations investigated by Chen and Dawson: nitrogen on an oxygen site, the nitrogen interstitial and nitrogen on a titanium site. From the densities of state, we see that the negative defects produce non spin-polarised electronic configurations while other charge states produce spin-polarised results. Reprinted with permission from Chen and Dawson [140]. Copyright 2015 American Chemical Society.



(a) N_O substitutional defect

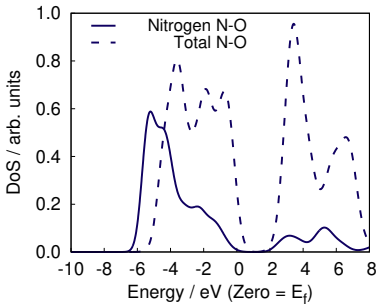


(b) N_{Ti} substitutional defect

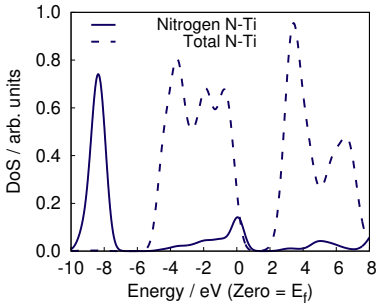


(c) N interstitial defect

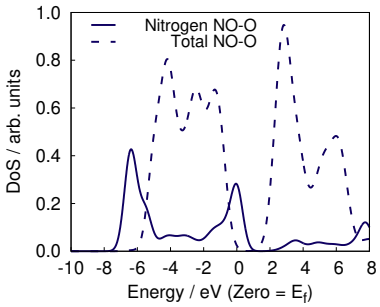
Figure 5.2: Optimised geometries (PBE level) of three possible nitrogen defects in bulk anatase.



(a) N_O substitutional defect



(b) N_{Ti} substitutional defect



(c) N interstitial defect

Figure 5.3: Density of states for anatase with nitrogen defect. Total DoS for system in dashed line - scaled down by a factor of 400. Solid line is contribution from nitrogen atom.

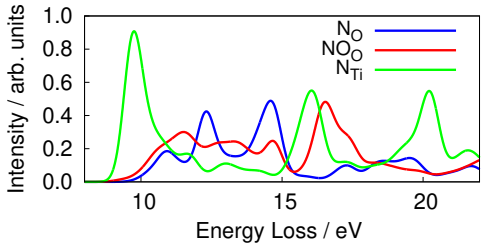


Figure 5.4: N K edges for the three different nitrogen defects in bulk anatase. No core hole was used here. The nitrogen titanium substitutional defect gives a spectrum which if most distinctive, this may be because the large difference in electronegativity between nitrogen and titanium results in the substitution producing a large change in the local electronic structure.

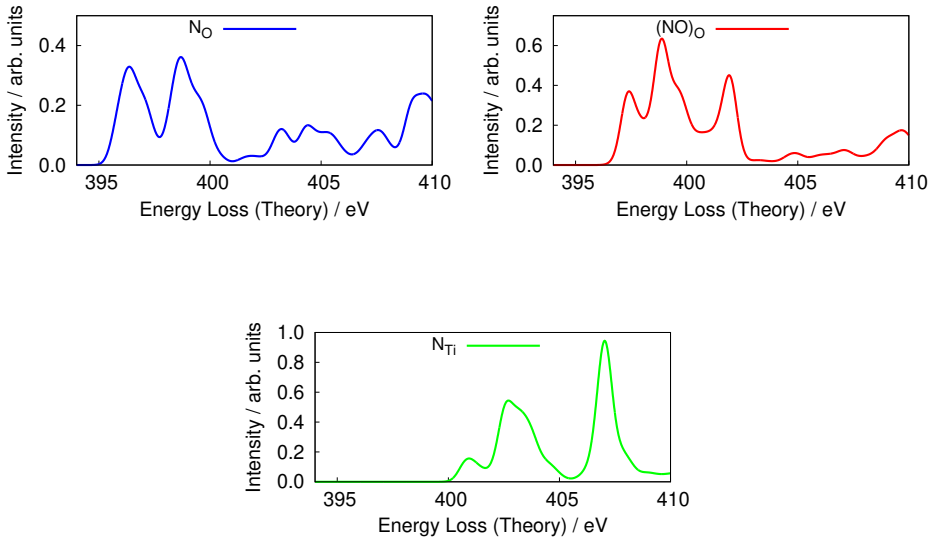


Figure 5.5: N K edges for three different nitrogen defects in bulk anatase. A full core hole was used here. Each defect clearly has a unique signature. The titanium site substitutional has a large peak arround 7 eV above the edge onset. This differs from the oxygen site substituational and interstitial which both show a relatively low instesity in this region. The key distinguishing feature between the oxygen site substitutional and the interstitial is the number of peaks and thier relative intensity. The oxygen site substitutional has a double peak with a spacing of arround 3 eV and a 1:1 ratio, while the interstitial has a tripple peak with spacings of 2 eV and 3 eV and an approximate 1:2:1 ratio. All three plots have the same x axis to allow easy comparison of alignment. The theoretical edge offsets were produced using the appoach of Mizoguchi [106].

Chapter 6

Surface Defects in Anatase: Formation Energy

Many studies have been conducted on the anatase (101) surface, covering topics including small molecule adsorbtion, heteroatom doping and intrinsic defects [119, 137, 138, 144–150]. The plethora of existing studies looking at oxygen vacancies may lead an observer to think that the field is largely well understood but as we will show, these studies tend to suffer from identifiable shortcomings. Firstly, the limitations of conventional DFT codes has often necessitated the use of rather small simulation cells, which exacerbates finite size effects (see Sections 3.5 and 6.3). Secondly, and as a result of these finite size errors, many authors have chosen to treat only the neutral oxygen vacancy. In this chapter, through a combination of large supercells enabled by the ONETEP code and a finite size correction scheme chosen for its ability to handle complex surfaces, the objective is to make robust conclusions about defects in and near the anatase surface.

The challenges of finite size effects will be discussed and demonstrated in section 6.3. Following this, several schemes for correcting the electrostatic component of the finite size error in formation energies will be examined and the sutability of the method ultimately seletected estabilished. In section 6.6

some of the considerations in implementing the chosen method are discussed, such as how the sensitivity to its adjustable parameters was measured and how those parameters were selected.

6.1 Prior Work

The many technological applications of anatase, both present and envisaged, have made its defects an attractive target for computational study. DFT studies have been conducted using plane wave codes [151]. The DFT+U methodology has also been used, for example, by Morgan and Watson [152]. Depending on the choice of methodology, there is significant variation in the observed electronic properties of some defect species. For example oxygen vacancies in Na-Phattalung *et al.* [151] generate no states in the band gap, whereas Morgan and Watson find that they do [152].

The formation of defects at or near the anatase (101) surface is of particular interest, as this facet has a low surface energy and thus makes up the majority of the surface on an anatase nanoparticle [153]. Despite the difficulties outlined by Ha and Alexandrov [154], some consensus has developed regarding the properties of defects in this surface. Prime amongst these established properties is the result, also seen in this work, that oxygen vacancies on the surface itself tend to be unstable with respect to the subsurface layer [119, 148–150, 155–157]. This result has been obtained by workers using both GGA (this work, and Cheng and Selloni [119, 157]) and screened exchange [120], as well as being corroborated by experimental observations [149, 150]. It is worth noting that Haa and Alexandrov [154] obtain this result only in the case of closed shell and spin-polarised PBE-DFT without U or D terms, whereas the use of Hubbard U, dispersion, or hybrid functionals leads to the surface site becoming favourable. Haa and Alexandrov's work shows that the choice of functional can have a profound impact on the results obtained for simulations of the anatase surface. It is

uncertain whether these parameters are actually providing a more reliable model of reality, however, as scanning tunnelling microscopy seems to indicate that defects in the surface itself are rare, at least in freshly cleaved (101) surfaces [149, 150].

There are clearly deficiencies in the use of semi-local functionals without additions such as Hubbard U for the study of anatase defects, particularly in cases where states in or above the band gap are occupied: for example, anatase samples containing defects have been observed experimentally to have magnetic properties [158, 159]. Yang *et al.* have shown that these can be predicted in simulations by the inclusion of an appropriate U term [160]. Furthermore, when dealing with some defects such as the neutral oxygen vacancy, semi-local functionals may lead to de-localisation of electrons. This behaviour is not seen at higher levels of theory which instead give electrons localised on titanium sites [152].

Calculations using hybrid functionals in titanium oxides have been undertaken by Finnazzi *et al.* [161], who determined that only hybrid functionals or GGA+U were able to reproduce experimentally observed gap states. Further work by Janotti *et al.* and Mattioli *et al.* established oxygen vacancies to be shallow donors [162, 163], suggesting that the +2 state will dominate in most doping regimes. Despite its apparent stability, the +2 state has seen less interest, possibly owing to difficulties in controlling finite size effects. This defect will thus be the main focus of this chapter.

6.2 Simulation details

Initial calculations were conducted using the same parameters as in Section 4.3.1: NGWF radii of 10 Bohr, a cutoff energy of 800eV and the JTH PAW datasets [108]. Once again, the PBE functional [37] is used. As for the doubly charged defect, there are no occupied states above the top of the valence band and a semi local functional can be expected to perform well. A

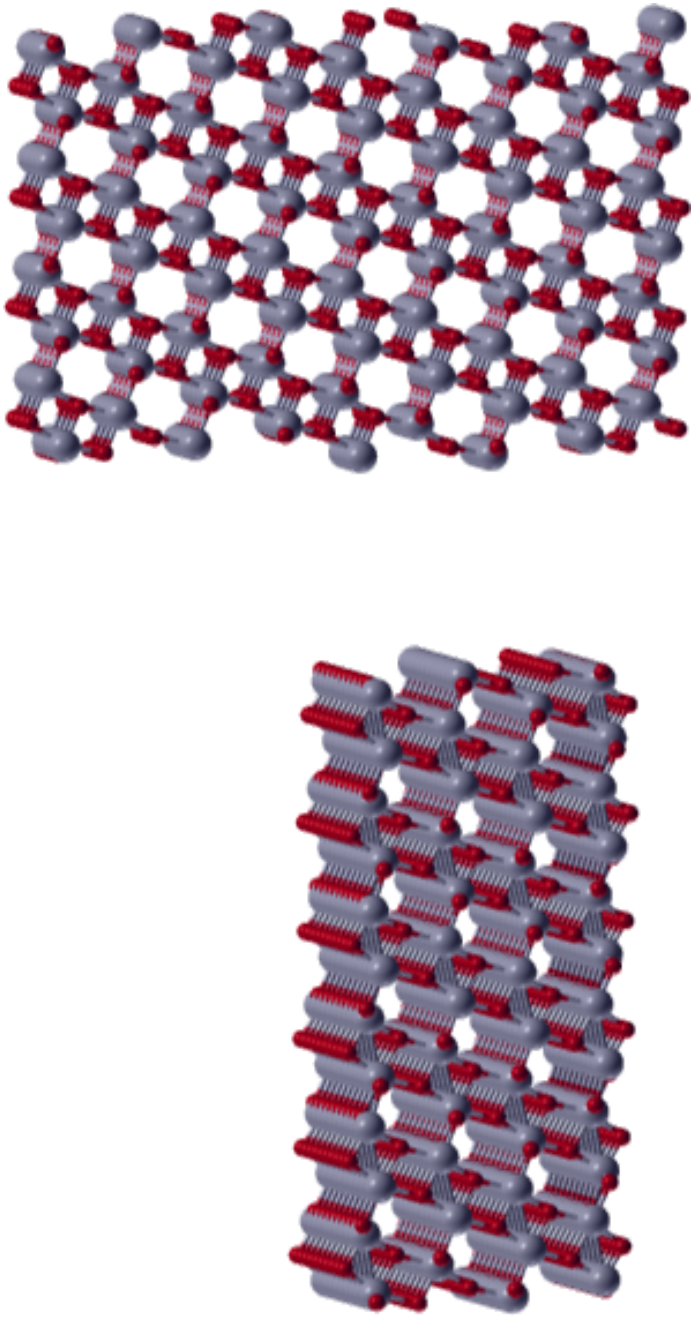


Figure 6.1: The 1440 atom 'large' cell (left) and 864 atom 'thick' cell (right).

series of cells with defects based on the 720 atom anatase (101) surface cell are shown in Figure 4.1. These cells were subjected to geometry optimisation with a tolerance of 0.10 eV/Å. Most cells were optimised to better than this in practice, with some achieving a force tolerance of 5.1×10^{-2} eV/Å.

For comparison, a bulk cell (384 atoms) based on the same primitive cell was also produced. Finally, to enable tests of convergence of defect formation energies, two larger surface slabs were also produced. One slab was based on a doubling of the 720 atom slab along the b axis and contained 1440 atoms. The second was a thicker slab, intended for testing convergence of formation energy with depth into the sample. This cell contained 864 atoms. These final two cells are depicted in Figure 6.1.

6.2.1 Hybrid Functionals

Calculations with the PBE functional for the +1 and neutral charge states of the oxygen vacancy showed a tendency for charge to delocalise from the defect. To investigate the role of self interaction errors in this effect, calculations were conducted using the HSE06 hybrid functional. Due to the increased computational cost of this functional compared to PBE, the calculations were conducted in smaller cells (48 or 108 atoms) and the CASTEP code was used. Initial geometry optimisation was conducted using the PBE functional, as this saved a considerable amount of time. The optimised PBE geometries were then used for the ground state hybrid calculations. Limitations in the implementation of hybrid functionals in CASTEP meant norm conserving pseudopotentials from the OPIUM database of Rappe and Bennett [164] were used.

$\|\vec{a}\|$ / Å	$\|\vec{b}\|$ /Å	Vacuum gap / Å	Uncorrected E_f / eV	Corrected E_f / eV
21	11	13	2.89	3.40
21	22	13	3.02	3.58
42	11	13	3.25	3.60
42	22	13	3.26	3.75
21	11	24	3.79	3.56
21	22	24	3.45	3.65
42	11	24	3.67	3.66
42	22	24	3.40	3.70
21	11	34	4.76	3.56
21	11	39	5.34	3.59
63	33	13	3.45	3.83

Table 6.1: Dimensions of test cells used, along with corrected and uncorrected (see section 6.9) formation energies. $\vec{a}$ and $\vec{b}$ are the in plane lattice vectors.

6.3 Finite Size Effects

To study the importance of finite size effects in the anatase surface system, an oxygen vacancy was simulated in cells with a variety of in plane sizes and vacuum gaps. The cell parameters are shown in Table 6.1. The results of this test are shown in Figure 6.2. Without accounting for finite size effects, the results show a great deal of spread. Note in particular the strong divergence with increasing vacuum gap. This is clearly inadequate for predictive purposes, so in the following sections we address the main origins of finite size errors and propose and implement appropriate solutions.

6.3.1 Defect Bands

If an isolated defect has localised wavefunctions associated with it, the finite extent of these wavefunctions must be taken into account when selecting a

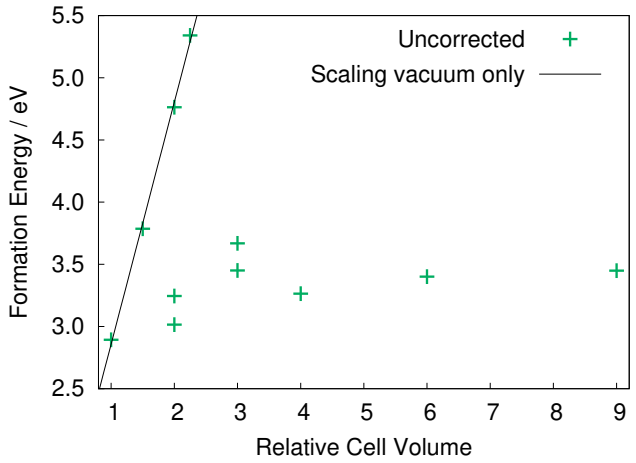


Figure 6.2: Formation energy for a +2 oxygen vacancy defect in anatase surfaces with various in plane sizes and vacuum gaps. The line is to guide the eye in showing the divergence in the results for a series of cells.

supercell size. If the supercell is too small, the tails of defect wavefunctions may overlap with those of the periodic images of the defect. This leads to the formation of a defect band with finite bandwidth, which, depending on the position of the defect state in the gap, may then hybridise with the conduction or valence bands. The formation of defect bands can depend sensitively on the level of theory used, as the position of defect levels in the gap can vary greatly with the choice of functional [154]. If defect states are well localised, then even a small cell may be sufficient to prevent interaction. As defect levels move closer to band edges, there is increasing de-localisation, which necessitates larger cells to prevent finite size errors resulting from the formation of spurious bands. Examination of the band structure of the simulation cell can enable the band width of any defect states to be established. A high bandwidth may indicate an insufficiently small simulation cell.

6.3.2 Elastic

The introduction of a defect into a crystal disrupts the bonding of neighbouring atoms. This disruption may be very localised in the case of some defects in rigid crystal lattices, such as NaCl. In the case of TMOs, however, there can be significant long ranged relaxations. Surfaces produce additional complexity, as they provide a direction without material into which the crystal can relax more extensively.

Elastic effects are very material specific and as such, no single scheme for correcting them exists. The best option is to use sufficiently large cells such that there exists a region between defects where there is a sufficiently small change in atomic coordinates from the expected bulk values. This can be checked either directly, or by monitoring convergence of formation energy with supercell size. ONETEP is an ideal code here, as large simulation cells can be used with relative ease, allowing systematic investigation of convergence.

6.3.3 Electrostatic

In vacuum charges interact according to the $\frac{1}{r}$ Coulomb law. In insulators, this interaction is screened by the dielectric constant of the material, but its functional form is unchanged. This leads to extremely long-ranged interactions: doubling a cell in every dimension will increase the number of atoms eightfold but only halve the electrostatic interaction error. Even with linear-scaling DFT, good convergence with cell size can not be feasibly achieved simply by using ever-larger simulation cells.

Unlike elastic effects, the Coulomb interaction has been subject to extensive analytical work. This means it is possible to develop expressions for the expected strength of the interaction with distance, even in materials with non trivial dielectric tensors. For example, Turban *et al.* addressed the issue of dipole-dipole interactions in pentacene [165]. Once such expressions have been derived, it is then possible to use them to correct DFT formation ener-

gies from finite simulation cells to obtain estimates of formation energies for the isolated defect. Several of these schemes are reviewed in the next section.

6.4 Correction Schemes

To address the difficulties posed by long ranged electrostatic interactions, several schemes have been developed [166–168]. These are based on a number of different approaches, though most are derived from classical electrostatics.

6.4.1 Makov-Payne

One of the first works to address the issue was Makov and Payne [166]. Using classical electrostatic arguments, an expression for the electrostatic energy of a charged defect in a periodically repeated cubic supercell defined:

$$E = E_0 - \frac{q^2\alpha}{2L\epsilon} - \frac{2\pi qQ}{3L^3\epsilon} + \mathcal{O}(L^{-5})\,, \tag{6.1}$$

where q is the charge of the defect, α is the Madelung constant, L is the side length of the cell, ϵ is the bulk zero frequency dielectric constant of the material and Q is the radial moment of the defect charge distribution.

Some of the quantities used in Eq. 6.1 are not readily accessible from calculations: ϵ and Q cause particular difficulty. To address this, Eq. 6.1 is generally not used to correct the results of a single supercell calculation but instead several calculations at varying supercell sizes are conducted and the resulting energies are then fitted [169]. This has the additional advantage that an estimate of ϵ can also be obtained from ground state calculations.

The definition of Eq. 6.1 is restricted only to cubic cells in the presence of a scalar dielectric. Many systems of interest have neither of these properties. Several extensions to the original Makov-Payne scheme have been developed. As an example, Hine *et al.* [170] used an modified version of the Makov-Payne scheme where supercell energy was fitted not to supercell size but to the

Madelung potential of the charge distribution for that supercell. This allowed the use of simulation cells with differing aspect ratios, greatly increasing the number of data points available for extrapolation to infinite size.

Attempts were made early in this investigation to extend the fitting scheme of [170] to cells with an anisotropic and spatially-varying dielectric tensor, however this approach lacked a rigorous means of handling spatial variation, such as the presence of a surface, and was thus abandoned.

6.4.2 Freysoldt, Neugebauer and Van de Walle

To address some of the issues with the Makov-Payne and other schemes, Freysoldt *et al.* developed a method working directly with those quantities available within a DFT calculation [167]. In this approach the contributions to the total energy of a defect within a periodic lattice are divided into long-range and short-range contributions, E^{inter} and E^{intra}:

$$\begin{aligned} E_{\text{tot}} &= E^{\text{inter}} + E^{\text{intra}} \\ &= E_q^{\text{lat}} + q\Delta_{q/0}\,. \end{aligned} \tag{6.2}$$

The first term in the second line represents the long range interaction of the defect charge with its periodic images and the compensating background charge. The second term has the form of a potential alignment term and is written as the average over the cell of the short range part of the potential due to the defect charge (Eqs. 14 and 5 of Ref [167]):

$$\Delta_{q/0} = \frac{1}{\Omega}\int_{\Omega}\left(\tilde{V}_{q/0}(\mathbf{r}) - \tilde{V}_q^{\text{lr}}(\mathbf{r}) - C\right)\mathrm{d}^3\mathbf{r}\,. \tag{6.3}$$

The integrand here is the difference between the total periodic potential available from DFT and the long-range periodic potential derived from the Coulomb potential. The long-range periodic potential is produced by removing the $\mathbf{G} = 0$ term in the Fourier space representation of $\frac{1}{r}$.

The short-range component is produced by subtracting the long-range model potential from the potential obtained from DFT calculations. This

step allows the correctness of the method to be confirmed. Plotting a 1D sample of the short range potential should show it tending to a constant value (see Figure 2 of Ref. [167]). This constant value must be taken into account in the final energy calculation: it provides the potential alignment term C.

This scheme requires a charge distribution be specified. However, it is noted by the authors that if a sufficiently localised charge distribution is used then results are largely independent of the specific choice. Gaussian and point charge distributions are suggested.

This scheme does not readily treat anisotropic materials or surfaces as formulated because it uses a single scalar dielectric constant. As a result, it could not be used for this study.

6.4.3 Komsa-Pasquarello

Rather than deriving potentials from DFT calculations, the method of Komsa and Pasquarello [168] directly computes potentials, and thus electrostatic energies, for test charge distributions in a continuum model system with a dielectric profile which is fitted to the real (DFT) system of interest. Within the model system, the energy of an isolated defect charge distribution is also calculated. This is done either by direct analytic methods for simple defect geometries, such as a point charge in a uniform anisotropic dielectric, or by extrapolation of periodic systems with increasing size for more complex cases, such as a Gaussian charge distribution near the surface of an anisotropic dielectric slab.

This scheme has the great advantage that all potential calculations take place within the model system and can be tackled using FFT based methods. Furthermore, the choice of geometry for the system is not limited to bulk solids. A crucial advantage for the study of anatase is that the Komsa-Pasquarello method can readily be extended to handle systems with an anisotropic dielectric.

A key quantity in the Komsa-Pasquarello method is the dielectric model used. In all cases an estimate for the bulk zero frequency dielectric constant of the material of interest is required. For surfaces the roll off from bulk the vacuum values of ϵ is accomplished using an error function profile. The thickness of the slab and error function characteristic length then become adjustable parameters. Two possible routes could then be followed: these parameters could be fitted to minimise the spread of corrected results, or the parameters can be chosen to correctly reproduce the x-y plane averaged electrostatic potential extracted from DFT calculations (Here the slab normal is taken to point along z). The first option is unsatisfying and could yield fitted parameters which do not make physical sense. The second option is therefore preferred. Once the model periodic and aperiodic energies have been obtained, the DFT energy can then be corrected by applying a correction to the total energies:

$$E_{\text{cor}} = E_{\text{isolated}} - E_{\text{periodic}} + q\Delta V \,. \tag{6.4}$$

The final term accounts for alignment between the DFT and model potentials. ΔV is the difference in potentials far from the defect, and q is the defect charge.

This scheme has been tested by its authors on simple surfaces, namely NaCl and hydrogen passivated GaAs. These results motivated the decision to implement a version of this finite-size correction method and attempt to use it on the anatase surface system.

6.5 Computing a Formation Energy

The typical output of a DFT calculation is the total energy of the system. This is denoted here by E_{sys}. To go from this energy to a formation energy E_f, the immediate idea is to take a total energy difference between a perfect

and defective cell

$$E_f = E_{\text{defect}} - E_{\text{perfect}} \,. \tag{6.5}$$

Unfortunately, for a general defect the number of atoms in the perfect and defective system may not be the same. To address this problem, the chemical potentials μ_x of the species lost and gained, including any electrons associated with a charged defect, must be taken into account. A common expression for this is by Zhang and Northrup [171]:

$$E_f = E_{\text{defect}} - E_{\text{perfect}} + \mu_x + qE_{\text{Fermi}} \,. \tag{6.6}$$

In this section, the method of Finnis *et al.* [172] is used. This scheme is particularly useful for oxygen vacancies as other methods of computing the oxygen chemical potential (say, from an isolated molecule in a box) are rather unreliable. Instead, chemical potentials are fixed via a Born–Haber cycle which is underpinned by tabulated experimental Gibbs energies.

6.6 Implementating the Komsa-Pasquarello Method

To enable rapid prototyping and development, the python programming language was used to implement an FFT based Poisson solver for use with the Komsa-Pasquarello method. The poisson equation being solved may be written as

$$\vec{\nabla} \cdot \underline{\underline{\epsilon}} \cdot \vec{\nabla}\phi = -\rho \,. \tag{6.7}$$

Here $\underline{\underline{\epsilon}}$ is the dielectric tensor, ϕ is the potential and ρ the charge density. As this is a vector equation $\cdot$ is used to denote matrix multiplication.

The code used the scipy and numpy libraries for acceleration. The resulting code was tested using the NaCl system described in [168] and also against

analytical results for a Gaussian charge distribution in a uniform dielectric environment.

Once the correctness of the code had been established, extensions as described in the supplementary information of [168] were implemented. These enabled the use of a diagonal anisotropic dielectric, which allows for a very close approximation to the dielectric tensor of the anatase surfaces under investigation, whose off diagonal terms are small.

6.7 Electrostatic Energy In Periodic Boundary Conditions

In a periodic environment, some care must be taken when discussing the energy of a charge distribution. It is clear that if the periodic unit cell has a net charge then the overall energy of an infinite lattice will be divergent. To overcome this problem, a cancelling background charge is generally assumed to exist in any 'charged' DFT calculations. This background is not introduced explicitly, but rather by omitting divergent terms from the sums used to compute overall electrostatic energy. In the presence of a dielectric, the expression for the energy, E, of a charge distribution is

$$E = \frac{1}{2}\int \mathbf{D}\cdot\mathbf{E}\,\mathrm{d}\mathbf{r}\,. \tag{6.8}$$

Where $\mathbf{E}$ and $\mathbf{D}$ are the electric field and electric displacement of classical electrostatics respectively. Using the expressions for $\mathbf{D}$ and $\mathbf{E}$ in terms of the potential, V, and charge distribution, ρ, we obtain:

$$\begin{aligned} E &= \frac{1}{2}\int \mathbf{D}\cdot\nabla V\mathrm{d}\mathbf{r} \\ &= \frac{1}{2}\int \nabla\cdot(\mathbf{D}V)\mathrm{d}\mathbf{r} + \frac{1}{2}\int(\nabla\cdot\mathbf{D})V\,\mathrm{d}\mathbf{r} \\ &= \frac{1}{2}\int \nabla\cdot(\mathbf{D}V)\mathrm{d}\mathbf{r} + \frac{1}{2}\int \rho V\,\mathrm{d}\mathbf{r} \end{aligned} \tag{6.9}$$

Finally, using the periodic boundary conditions, we argue that the first term is identically zero, leaving the final expression

$$E = \frac{1}{2} \int \rho V \mathrm{d}\mathbf{r} \,. \tag{6.10}$$

6.8 Model System

6.8.1 The Dielectric Constant

In order to create the model system, values for the components of the zero frequency dielectric tensor are required. The zero frequency dielectric tensor describes the response of a material to a static (non time varying) electric field. There are three possible options for obtaining this tensor. Firstly, the experimental value could be used. This has the disadvantage that it will not be consistent with the DFT value, which is what is responsible for screening the defect charge in a simulation. Furthermore, in the case of anatase this value is not well-characterised experimentally. Secondly, the dielectric constant could be determined by applying a constant electric field to a slab simulation and observing how the potential in the material varies. This method has the advantage of providing information about the local dielectric environment of the surface, which can lead to the production of more accurate model systems. Unfortunately, attempts to use this method for the anatase system suffered from convergence problems, with applied fields either being too high and pulling electron density out of the slab, or being too low to produce a detectable effect. The final option is to use density functional perturbation theory to compute the dielectric constant. This method computes the full dielectric tensor for a material in the bulk. Tensor transformation can be used to express the dielectric tensor in terms of a different set of lattice vectors from that used in the initial DFPT simulation. This was necessary in the case of anatase, as the axes of the surface slab model were not coincident with those of the bulk system. Using this final method,

the dielectric tensor of anatase was computed to be

$$\begin{pmatrix} 33.8 & 0 & 0 \\ 0 & 33.8 & 0 \\ 0 & 0 & 22.24 \end{pmatrix} . \tag{6.11}$$

Applying an appropriate tensor transformation gave the tensor expressed in terms of the lattice vectors of the surface cell as

$$\begin{pmatrix} 30.9 & 0 & 5.0 \\ 0 & 33.8 & 0 \\ 5.0 & 0 & 25.14 \end{pmatrix} . \tag{6.12}$$

Literature values for the zero frequency component in the (101) direction include 35.0 [173], 39.7 [174], 41.0 [175] and 41.9 [176]. These are all within 36% of the value obtained in this work. These large errors are to be expected as dielectric tensor values involve the second derivative of total energy and are thus much more sensitive to the computational methods used than the total energy itself. The dielectric environment of the slab system was expressed in the same terms as in [168] using an error function of the form

$$\epsilon(z) = \epsilon_1 + \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{z - \frac{t}{2} - d}{\beta}\right)\mathrm{erf}\left(-\frac{z + \frac{t}{2} + d}{\beta}\right)\right)(\epsilon_2 - \epsilon_1) \ . \tag{6.13}$$

The parameters were chosen to maximise agreement between potentials computed using the model system and those from the DFT simulations. The meaning of these parameters in illustrated in Figure 6.3.

6.8.2 Charge Distribution

In order to generate model potentials, the Poisson solver requires a test charge distribution. It should be immediately pointed out that while the electrostatic energy of a test distribution depends greatly on its shape, if the *difference* in energies between an isolated and a periodic distribution with

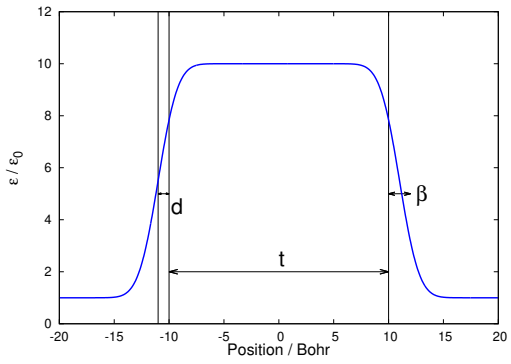


Figure 6.3: The dielectric parameters used. t is the thickness of the slab, d is the distance above the slab surface of the center of the roll off, β is the width of the roll off.

the same shape is considered, then providing the distribution is localised and the result becomes quite insensitive to the shape. Table 6.2 shows this in the context of an anatase slab. Variations of 50% in the value of σ have little impact on the value of the correction energy obtained. For example, the surface defect (index 1) shows the greatest variation in correction energy with distribution width, σ, with a change of 3.4% on varying σ between 0.7 and 1.0 bohr, while keeping other model parameters constant.

As computed correction energies are not overly sensitive to the test charge distribution used, a simple Gaussian profile was chosen. This captures the qualitative properties of the charge distribution seen in DFT results (localisation and approximate spherical symmetry) while having a functional form admitting analytic results. The profile adopted is

$$\rho(r) = \frac{q}{\sigma^3(2\pi)^{\frac{3}{2}}} e^{-\frac{r^2}{2\sigma^2}}, \tag{6.14}$$

where σ is the distribution width and q is the total charge.

The charge distribution used for modelling the anatase system was produced by fitting a Gaussian to the difference in pseudoatomic charge between the perfect and defective cells. This produced a distribution with a σ value

of 1.6 Bohr.

Other methods of obtaining a model charge distribution were also considered. These included: fitting the distribution radius to best generate the potential seen in DFT calculations; and fitting the charge distribution to the square of the defect wavefunction. The former was dismissed as being too *post hoc* and the latter suffered from the delocalised nature of the defect wavefunction at the GGA level of theory.

Except in the case of defects at the surface of a slab, the energy correction term is relatively insensitive to the shape of the charge distribution used. This is because the correction depends on the long ranged electrostatics of the system where, per Gauss' law, the electrostatic potential varies little from that of a point charge.

6.8.3 Sensitivity to Parameters

To ascertain the magnitude of the errors in the final energy corrections introduced by changes in the parameters for the dielectric profile, a series of model simulations were conducted using all combinations of parameters from the sets $t \in \{12, 13, 14\}, d \in \{0.5, 0.75, 1.0\}, \beta \in \{0.6, 0.8, 1.0\}, \sigma \in \{0.71, 0.82, 1.0\}$. These values were selected such that the middle value in each set is approximately the optimal value. Table 6.2 shows the results of this study. It is clear that the choice of parameters for the dielectric profile can have a significant impact on the correction energy computed. It must be stressed, however, that the range of parameters used here vary by a factor of two for d and almost the same for β. The single most important value to get correct is the thickness of the slab, but as this can be directly measured from the atomic geometry, it is also the simplest value to get correct. Values of d and β are harder to ascertain but have a proportionally lower impact on the result.

The impact of the spacing of the simulation grid was also investigated. It was found that although the absolute energies computed converged slowly

Defect Index	t	d	β	σ
1	13.4	6.93	8.57	3.44
2	17.8	7.60	8.53	0.535
3	18.2	7.63	8.35	0.028
4	18.5	7.65	8.17	0.007
5	18.8	7.68	8.00	0.012
6	19.0	7.70	7.89	0.015
7	19.1	7.71	7.82	0.017

Table 6.2: For each defect index a correction energy was computed using models with each possible combination of the parameters from the sets in section 6.8.3. The results here show the spread of computed correction energies when only one parameter is allowed to change while the others are kept constant. For example to produce the first entry of 13.4% the values of d, β and σ were fixed at 0.75, 0.6 and 0.82 respectively and correction energies were computed with t values of 12 and 14. The final result is the difference in the two energies computed, expressed as a percentage of the smaller value. Fixed values were $t = 13.0\,, d = 0.75\,, \beta = 0.8\,, \sigma = 0.75\,.$ Ranges were $t = 12.0 \ldots 14.0\,, d = 0.5 \ldots 1.0\,, \beta = 0.6 \ldots 1.0\,, \sigma = 0.71 \ldots 1.0\,.$

with reducing grid spacing (equivalent to increasing the cutoff energy in a plane wave code), the difference between periodic and extrapolated energies converged more rapidly. Since it is this latter quantity which is required to correct formation energies, this was the quantity which was converged.

6.8.4 Choice of the Targeted Dielectric Environment

When computing the electrostatic energy of an isolated defect, it is necessary to consider what is meant by the isolated case. Following the example of Komsa and Pasquarello, these are *slab*, *semi-infinite* and *bulk*. The slab limit is that of a single defect in a slab of finite thickness in the z direction and infinite extent in the $x - y$ plane. The slab is isolated in the z direction. The physical situation represented by this model is a defect in a suspended membrane of material. The semi-infinite case is that of a defect in a block of material which occupies the entire half space $z < 0$, with vacuum in the region $z > 0$. This model represents a defect near the surface of large piece of material. Finally, for the bulk like case, this limit is the defect immersed in an infinite bulk of the material. The physical system here is a defect near the middle of a large sample of material. This limit can only be usefully applied to a defect some distance from the surface.

In this chapter, the final two limits are of particular interest. As a defect is placed deeper and deeper into a slab cell, these two limits should converge to the same result, as the impact of the surface is attenuated by increasing distance.

6.9 Tests Of The Correction Scheme

As discussed in Section 6.6, the implementation was tested against data from Komsa and Pasquarello [168] and analytic results. In the case of the latter, the Poisson equation was solved for a Gaussian charge distribution in

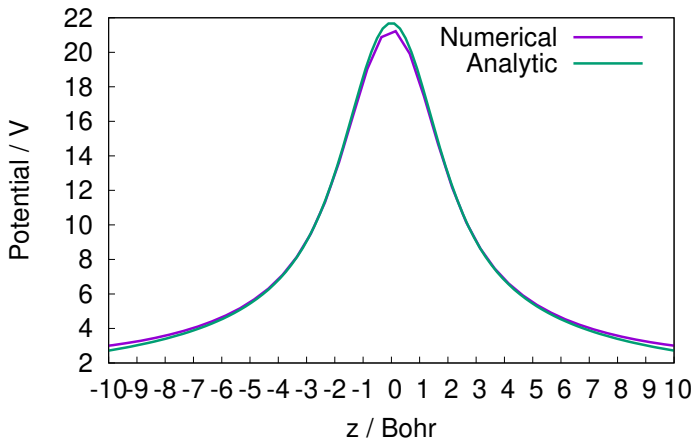


Figure 6.4: A comparison of analytic and numerical solutions for the Poisson equation for a Gaussian charge distribution immersed in an isotropic dielectric of relative strength 1. The Numerical calculation was performed in a cell with sides 30 Bohr long, and an FFT grid spacing of 0.5 Bohr. The test charge distribution had $\sigma = 1$.

a uniform dielectric. The resulting potential can be shown to be

$$V(r) = \frac{\mathrm{erf}(\frac{r}{\sigma\sqrt{2}})}{\epsilon r}\,, \tag{6.15}$$

so a direct comparison can be made.

A plot of the potential derived from numerical simulation is shown in Figure 6.4 along with the analytic result indicating good agreement across most of the range of z, with some regions showing greater discrepancy. Towards the edge of the cell, the periodic boundary conditions manifest themselves in a reduction of the gradient of the numerical solution. At the centre of the cell, the numerical solution does a poor job of representing the analytic solution. This is due to the limited size of the basis set which reduces the accuracy with which rapidly changing features can be represented. It is possible to go a step further and derive an analytic expression for the electrostatic energy

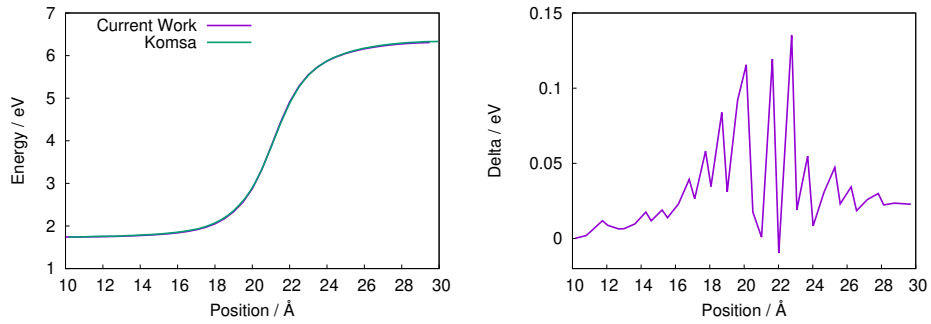


Figure 6.5: Comparison between results of Komsa and Pasquarello for the electrostatic energy of a charge distribution near a NaCl surface (at 20Å here). The parameters used were $t = 22.48$, $d = 2.3$, $\beta = 2.0$, $\sigma = 3.16$, $\epsilon_1 = \epsilon_2 = \epsilon_3 = 2.45$.

of an isolated Gaussian charge distribution in an isotropic dielectric. By substituting the potential in Eq. 6.15 and the charge distribution Eq. 6.14 into Eq. 6.10, we obtain:

$$E = \frac{1}{2\sigma\epsilon\sqrt{\pi}}\,. \tag{6.16}$$

For a Gaussian of width σ in an isotropic dielectric of relative strength ϵ the energy.

Data for NaCl were extracted from [168] and are shown plotted along with the results from the python implementation in Figure 6.5. Once again, excellent agreement is seen.

As discussed in Section 6.9.1, it is possible to construct analytic expressions for the electrostatic energy of a charge distribution in an arbitrary uniform dielectric. Using these results, the implementation of the Poisson solver was tested and Excellent agreement was seen.

6.9.1 Extrapolation

One method used to compute the formation energy of isolated defects is that of extrapolating the periodic formation energy to systems with increasingly

large cell sizes. This method has both advantages and disadvantages. In terms of code complexity, it is very simple with only a few further lines being needed to implement. As a disadvantage, this method requires rather large cells due to the L^{-1} convergence of energy with cell size. The use of large cells leads to heavy computational requirements, as significant memory is needed to store the Fourier representation of the potential, and the matrix operations used by the Poisson solver scale unfavourably with the number of grid points. Furthermore, the infinite size extrapolation is obtained as a y-intercept from a fitting procedure, which has the potential to introduce errors.

Testing in Anisotropic Dielectrics

The application of this method to anisotropic dielectrics had not previously been demonstrated by Komsa, though if had been alluded to in the supplementary information of[168]. This verifying that the scheme worked correctly in anisotropic dielectrics was essential. To this end, a series of checks were performed in cases where a reference energy couldbe determined by analytic methods.

Gaussian charge in a uniform (but not necessarily isotropic) dielectric. In these situations, it is possible to obtain an analytic expression for the energy of the distribution. For an isotropic dielectric, this is a relatively simple task when Gauss' law is used. For an anisotropic dielectric, it is necessary to perform a coordinate transform to render the Laplacian in the Poisson equation isotropic. This transforms the test charge distribution from a spherical one to one with ellipsoidal symmetry. The electrostatic energy of such a distribution can be calculated using the method described in [177] and the results showed agreement with the extrapolation method of better

than 1%. The final expression for the energy of the distribution is

$$\begin{aligned} a &= \epsilon_1^{-\frac{1}{2}}(\epsilon_1\epsilon_2\epsilon_3)^{\frac{1}{6}}\,, \\ b &= \epsilon_2^{-\frac{1}{2}}(\epsilon_1\epsilon_2\epsilon_3)^{\frac{1}{6}}\,, \\ c &= \epsilon_3^{-\frac{1}{2}}(\epsilon_1\epsilon_2\epsilon_3)^{\frac{1}{6}}\,, \\ \nu_1 &= (1-\frac{a^2}{c^2})^{\frac{1}{2}}\,, \\ \nu_2 &= (1-\frac{b^2}{c^2})^{\frac{1}{2}}\,, \\ k &= \frac{\nu_2}{\nu_1}\,, \\ \sin(\phi) &= \nu_1\,, \\ \epsilon &= (\epsilon_1\epsilon_2\epsilon_3)^{\frac{1}{3}}\,, \\ E &= \frac{1}{c\nu_1}F(\phi,k)\frac{1}{2\sigma\epsilon^{\frac{5}{4}}\sqrt{\pi}}\,, \end{aligned} \tag{6.17}$$

where $F(\phi, k)$ is an incomplete elliptic integral of the first kind and a, b, c are defined in terms of the components of the dielectric tensor such that $c \geq b \geq a$.

6.9.2 Comparison with COFFEE Code

After development on the Poisson solver described in this work was underway, a paper describing a tool to achieve the same results was published by Naik and Jain [178]. Their tool is called CoFFEE. The Poisson solver was tested against the CoFFEE code to assess the correctness of the implementation. Figure 6.6 shows the potential along the z axis in the centre of a cell as calculated by both the present implementation and the CoFFEE code. It is clear that the two potentials are effectively identical. Computation of electrostatic energies yields discrepancies of approximately 15 meV, equivalent to 1%.

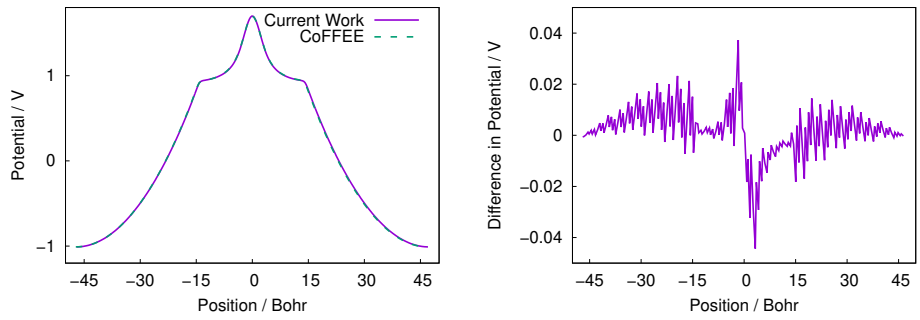


Figure 6.6: Electrostatic potential of a test charge distribution in the centre of a 14Å slab. The parameters controlling the roll-off were: $\sigma = 0.85$Å, $\beta = 0.40$Å. The simulation cell measured $30.6 \times 19.3 \times 49.9$ Angstrom. The FFT grid spacing was 0.16Å. The second figure shows difference between COFFEE code and current Poisson solver. Some of the difference in the right hand figure is due to alignment of the underlying FFT grid: the COFFEE developers use an odd number of grid points where in this work an even number is used.

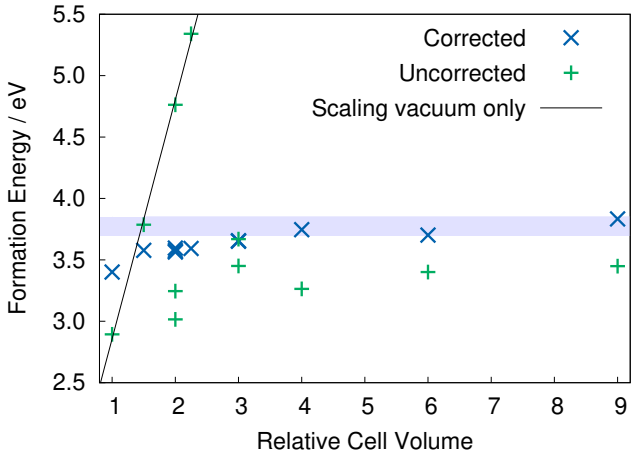


Figure 6.7: The results of applying the correction scheme to anatase surface slab cells with various shapes and sizes. The blue shading indicates the spread of corrected energies for the three largest cells. Once again, the line is to guide the eye in showing the divergence in the uncorrected energy for a series of cells.

6.9.3 Correction Of Test Surfaces

To demonstrate the power of the finite size correction scheme, we return to the results of Section 6.3. Applying a finite size correction to these results gives the formation energies shown in Figure 6.7. The most striking result here is the cancellation of the divergence in formation energy seen with increasing cell size. Also important is the reduction in the spread of the corrected energies compared to the uncorrected ones.

6.10 Results

6.10.1 Formation Energy With Depth

Satisfied that the finite size correction scheme was working, it was now possible to move onto the two primary objectives of this study: Firstly investigating the trend with depth of formation energies for V_O^{2+} at the anatase (101) surface. The second objective was to extend this by establishing if a trend with depth can be connected to the bulk value. This latter result would provide a measure of how far from the surface a defect must be before it can be considered 'bulk like'. In Figure 6.8, it is clear that when an appropriate dielectric environment is targeted, the formation energy of the defect converges rapidly towards the bulk value as depth into the slab (defect index) is increased.

It is immediately clear from Figure 6.8 that defects at position 5 are favoured with respect to the surface. The very low formation energy obtained with the 'bulk' correction scheme for defects at position 1 may be discounted as this results from treating a defect at the interface between two dielectrics as though it were immersed in only one. This result has been seen in previous studies [119, 120].

When defects with a lower charge state are examined, the same trend in formation energy with depth is seen as with the 2+ defect. However, overall formation energies are much higher and do not converge towards the formation energies obtained for bulk cells. The causes of this problem are dealt with in the next few sections.

6.10.2 Mulliken Population Analysis

To understand the degree of charge localisation (if any), Mulliken charges of atoms in the slab system were analysed. The system was broken into a series of layers (four for the 720 atom system) and the Mulliken charges

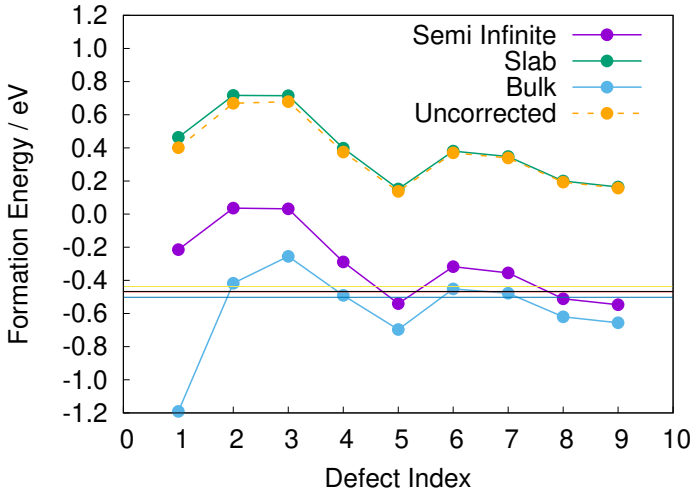


Figure 6.8: Trend in the formation energy ($E_f = 0$) of 2+ oxygen vacancies in the anatase 101 surface with depth. The three main lines indicate the three different ways of correcting the finite size electrostatic effects. The horizontal lines are a range of bulk defect formation energies computed using bulk cells with different sizes and aspect ratios. The orange dashed line gives the uncorrected formation energies - unsurprisingly this closely follows the slab corrected results.

Layer	Charge +2 / e	Charge +1 / e	Charge Neutral / e
1	0.503	0.097	-0.223
2	0.786	0.682	0.516
3	0.116	0.005	-0.173
4	0.600	0.214	-0.114

Table 6.3: The distribution of Mulliken charges between layers of an anatase surface model. The defect in each cell was positioned in the second layer.

for all atoms in each layer were summed. The results are shown in Table 6.3. It can be seen quite clearly that, as expected, the dielectric properties of the material lead to a partial screening of the defect charge with the net effect that a charge appears on the surface in the charged cell. The neutral defect cell exhibits behaviour consistent with charge delocalisation. The layer containing the defect (2) becomes positively charged as electron density migrates to the surfaces of the cell.

6.10.3 Charge Migration

One possible cause of the greatly different behaviour of the +1 and neutral defects, compared to the +2 defect, could be the fact that the +1 and neutral defects are expected to host one and two electrons respectively. It is possible that the levels associated with the +1 and neutral defects lie outside the band gap. If this is the case, then the states associated with the defect would not be localised and the electron density associated with them would be spread throughout the material. This would clearly hamper any attempt to treat the defect using a supercell scheme. The level position with respect to the band gap is sensitive to the choice of functional. Thus it is possible that GGA functionals may not reproduce the behaviour seen in the real material. Hybrid functionals are known to do a better job of treating electron localisation [161, 162]. If hybrid functionals also lead to a delocalised defect

state we may take this as further evidence that the defect state is delocalised in reality. Thus as an initial test using hybrid functional calculations was performed. This test also showed delocalisation of charge for the neutral defect.

6.10.4 Hybrid Functional results

Hybrid calculations show that the charge density associated with the neutral defect does not localise. To investigate if this was due to finding a spurious local minimum when lower energy spin polarised states are available, the simulation was re-run assigning an initial net spin to the simulation cell. This lead to the same result.

The lack of localisation my be due to missing polaronic effects, as the geometry of the cell was optimised using the PBE (GGA level of theory) functional. For example, Lambrecht suggests the importance of polarons in the study of defects [179] and Boonchun *et al.* base their decision to use hybrids on this [180]. Furthermore, the importance of polarons in the charge trasport properties of intrinsic antase are well known [181–183].

Can The Bulk Limit Be Reached?

Analysis of formation energy for defects deep in the thicker cells shows that in the case of the +2 defect, the bulk limit is reached fairly rapidly (defect index 4 to 6). In the case of the (partially) occupied defects, convergence to the bulk value is far more challenging. The use of hybrid functionals on a cell large enough to accommodate both a surface and a region of 'bulk like' material would be computationally unfeasible. DFT+U could be used as a semi-empirical method of achieving localisation. Even if DFT+U were to be used, there remains the outstanding issue of whether the correct state of the occupied defects is one where the electron states are localised. Work by Finazzi [161] using hybrid functionals suggests that DFT ground states with

some degree of delocalisation are very close in energy to those with localised Ti 3d electrons. This indicates that in the case of the +1 and neutral defects, the surface and bulk systems may behave in qualitatively different way.

6.11 Conclusion

In this chapter, the objective was to develop an understanding of defects near the anatase surface. This has been achieved: the doubly positive defect appears to behave as a localised charged defect and the neutral and singly charged cases exhibit more complex behaviour with electron density being delocalised.

The behaviour of electrons associated with +1 and neutral defects has been investigated using PBE and in the case of the neutral defect, also the hybrid HSE06 functional. Even at the hybrid level of theory, defect states are delocalised.

Across almost all of the gap, the +2 defect charge state is shown to be energetically favourable. This shows a problem with previous works which have tended to focus on the neutral defect, possibly to avoid the difficulties of correcting for the interaction between charged defect states. We have shown that with a method for addressing the problem of periodic images it is possible to correct image charge interaction errors in the anatase slab surface (Figure 6.2). The implementation of this scheme has been robustly tested against other published data. Large simulation cells accessible with the ONETEP code have been used to minimise the effects of long range elastic interactions.

Chapter 7

Line Defects in GaAs

The work underlying this chapter has been published in *ACS Nano Letters* as "Nonradiative Step Facets in Semiconductor Nanowires"
DOI: 10.1021/acs.nanolett.7b00123 [184].

Gallium arsenide is a frequently used semiconductor, with several properties which make it well suited to electronic applications, including a direct optical band gap with energies in the near infra red and high electron mobility. This latter property makes GaAs particularly appealing for radio frequency applications where it has maintained its position against silicon in transistors and diodes. This high switching speed has also attracted interest in using GaAs in new logic circuitry in next generation CPUs. These devices may no longer be manufactured using top-down methods, and thus an understanding of how defects in synthesised GaAs nanostructures behave is essential. Presently, whilst bulk GaAs is quite well understood, the properties of the nanostructured material are less well known, making both experimental and theoretical investigations essential. Here, a particular class of defect is investigated.

Defects and Recombination

When used for optical applications, attention must be given to the presence of any states in the material which can act as recombination centres. Deep level defect states are a common source, but anything which leads to a substantial disruption in the local band structure may have this effect. The optoelectronic properties of a material can be ruined by recombination with otherwise useful energy being dissipated thermally within the device, limiting both its efficiency and maximum operating power.

As a result of this, we sought to identify if the defect would produce any features in the band structure which could lead to either trapping or recombination.

7.1 Experimental Motivation

This work was conducted in collaboration with experimental groups at the University of Warwick lead by A. Sanchez and R. Beanland. After synthesising GaAs nanowires (with 5% phosphorus substituting for arsenic), scanning transmission electron microscopy was conducted. This revealed two classes of defect. The first of these is a twin defect, depicted in figure 7.1. As figure 7.1 shows, there is relatively little disruption in bonding in this case. The twin defect can be thought of as an interface between two regions related to each other by a 180 deg rotation about an axis perpendicular to the plane of the interface, with a {1,1,1} boundary. This is the $\Sigma 3(111)$ twin. The second class of defect is an addition to the twin, consisting of a step. In order for the step defect to be stable, the net Burgers vector must be zero or the defect would migrate to the surface of the nanowire and be expelled. It was proposed that stable step defects must adhere to a 'three monolayer' rule in order to be stable. This was based on experimental observations that only step defects containing three and six monolayers were present in the material. This class of defect is known as a $\Sigma 3(11\bar{2})$ step, as it is composed

of a section of $\Sigma 3(11\bar{2})$ boundary.

GaAs nanowire synthesis is typically achieved via a gas phase chemical deposition process, which may or may not be enhanced via the use of plasma. Wires grow from one end to the other, being deposited from a molten globule at the tip [185]. This layer by layer deposition process means that areas of a different orientation may end up embedded in the nanowire. Generally, stacking faults are thought to be expelled from a nanowire by image forces, but it can be shown that in certain cases no net force exists [184].

The nanowires were grown with a small amount of phosphorous, and the segregation of this dopant is discussed towards the end of this chapter.

7.1.1 GaAs in optoelectronics

As described in the introduction, the band properties of GaAs make it ideal for certain optoelectronic applications. Interest in using GaAs nanowire in similar roles is extensive. The main target application is solar energy [186, 187]. Light emitting diodes have also been proposed (GaAs in the bulk finds use in both conventional LEDs and LASERS) [188]. Other workers are interested in terahertz technologies and have developed both emitters and detectors [189, 190]. In all of these applications, device performance will be dictated by the presence of any unwanted defects.

7.2 Constructing Model Defects

To construct the defect model, a 12 atom orthogonal unit cell of zinc-blende structure GaAs was constructed. This cell was tiled in varying orientations to build up the model system. Figure 7.2 shows the structure of the basic cell and indicates how it was used to construct the defect models. The first defect studied was a simple twin. This is represented diagrammatically in figure 7.2b. The twin only has a minor impact on the coordination of

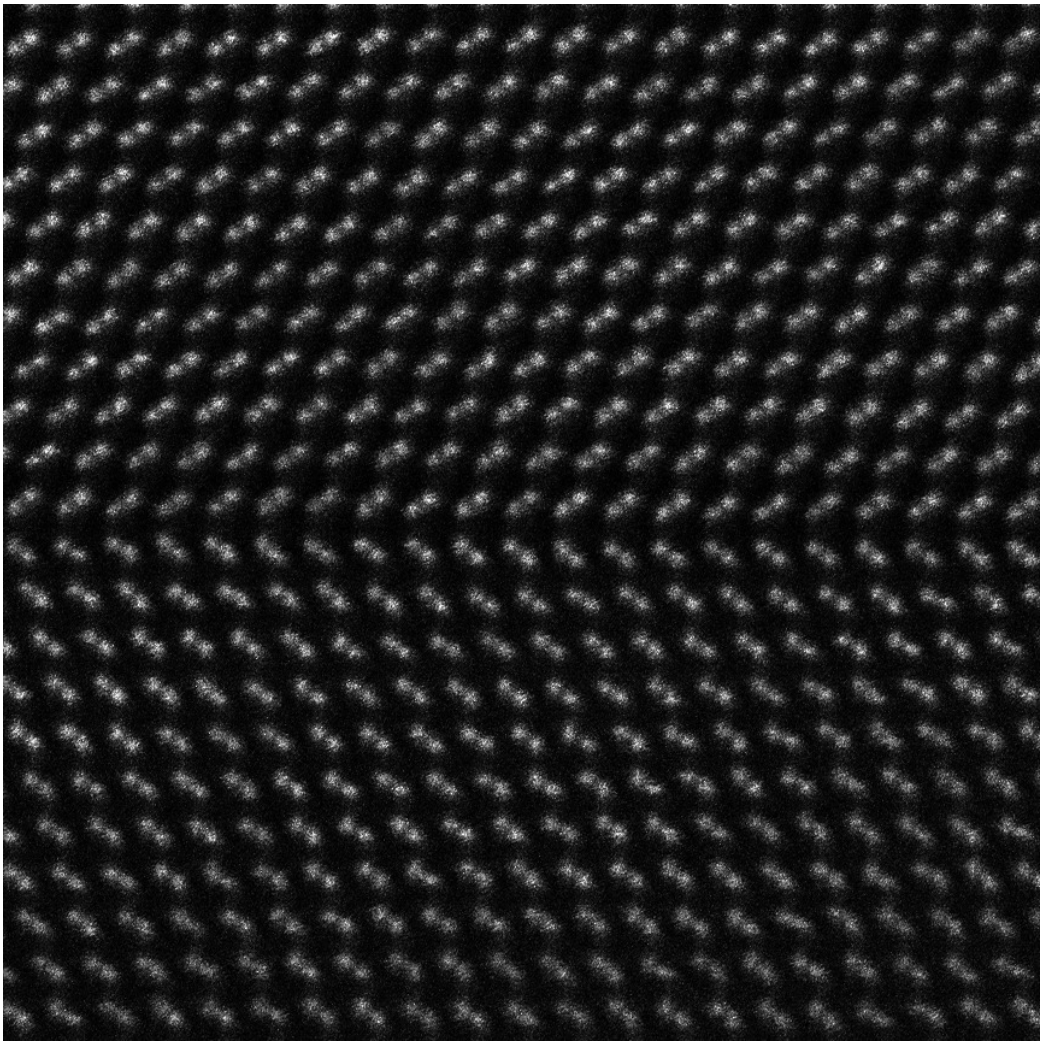


Figure 7.1: Electron micrograph of the simple twinning fault in GaAs.

atoms. The step edge defect has a much more profound impact on atomic coordination, as can be seen in figure 7.3.

While the defect model was produced by tiling a unit cell, TEM imaging of the defect core was also consulted to make sure the initial geometry used for geometry optimisation was a close as possible to that observed experimentally. The core of the defective structure is shown in Figure 7.3. Several features warrant discussion. Note that in this projection the bulk material, which is visible to the left and right, shows as a system of distorted hexagons. In the defect core, this structure is disrupted and a joint pentagon-heptagon system is seen. Crucially for later electronic properties, the defect core contains either an As-As or Ga-Ga bond, depending on the sense of the defect. The bond lengths in this case are found to vary somewhat from the values seen in bulk GaAs (2.47Å for the Ga-As bond in the bulk compared to 2.51Å for bonds making up the vertical sides of the pentagons in Figure 7.3), indicating weaker bonding.

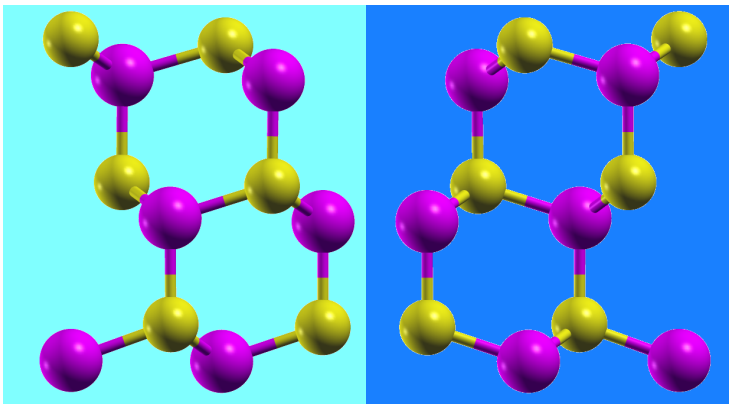
7.2.1 Creating Periodic Systems

The original defective system is observed in a nanowire and is therefore not periodic. In order to treat the defect system with a plane wave code, a periodic version of the defect needed to be constructed. This was accomplished by creating a large cell with two defects of the opposite 'sense'. This two-defect system could readily be modelled as a periodic system.

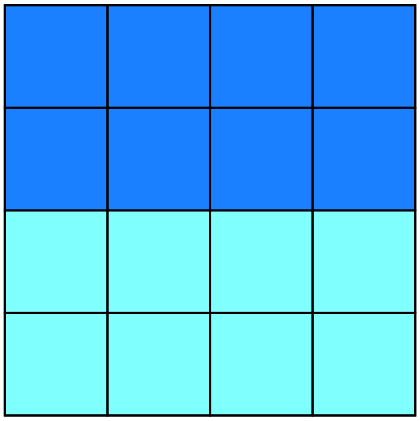
Two defect models were created, one containing a 3 layer step and a second containing a 6 layer step. This enabled testing to see if step height was a significant factor in electronic properties.

7.2.2 Simulation Methods

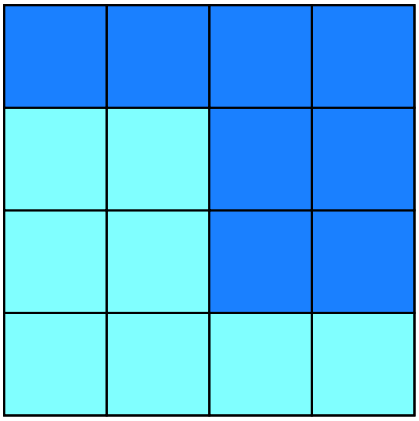
Several factors affected the choice of tool used for the simulation of the GaAs defect system. Firstly, in order to ensure reasonable isolation of the two



(a) The two base cells used to build the defect models



(b) The twin defect.



(c) The step defect.

Figure 7.2: The 12 atom cell used to build the defect model. Arsenic in yellow, gallium in purple. The second figure shows how the cell may be used to build up a simple twin defect. Light blue represents the original cell, dark blue represents a version rotated 180 degrees about the c (z) axis. The third figure shows how the more complex step defect is constructed.

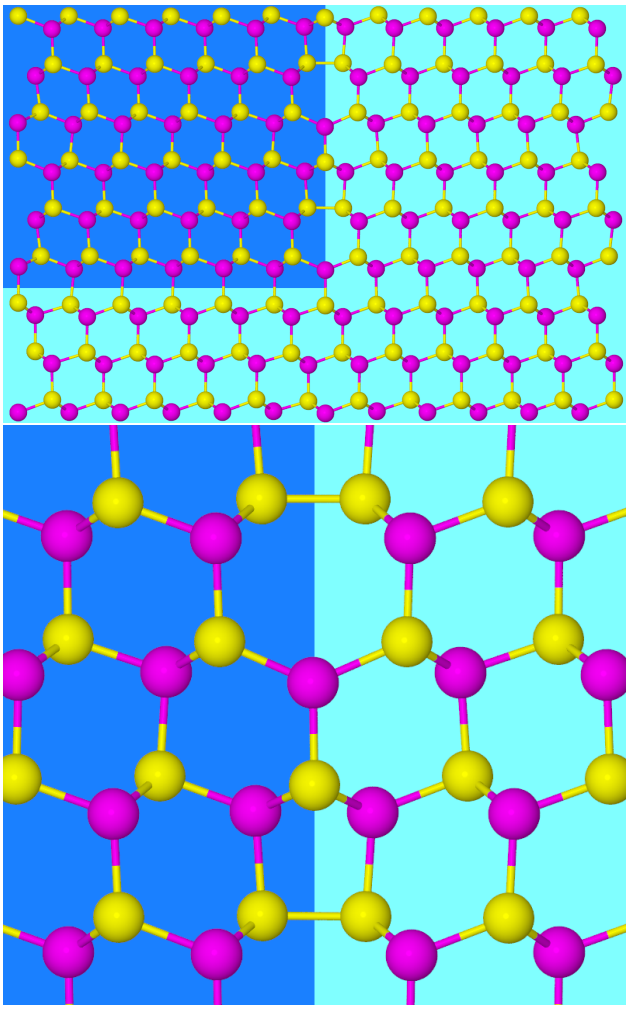


Figure 7.3: The first figure shows the structure of a six layer step defect. Note the heptagons and pentagons in the step region, along with the presence of As-As bonds. Arsenic in yellow, gallium in purple. The second figure shows a detail of the central area of the defect.

defects, they needed to be separated by a sufficient distance of bulk-like material. This necessitated a simulation cell which is large in the direction perpendicular to the defect axis. The defects also were assumed to show no disorder along the c axis (the axis into the page in figures 7.1 and 7.3). This meant that the simulation cell could be very thin in this direction. Along this direction, k-point sampling could be used. The k-point grid used was $1 \times 3 \times 1$. A plane wave cutoff of 700eV was used, along with CASTEP's *on the fly* pseudopotentials. Structures were relaxed using the BFGS algorithm as implemented in CASTEP.

7.3 Densities of States

The relaxed structures were used to produce densities of states for the system as a whole. The band gap was found to be effectively zero, reduced from 0.5eV in the case of the perfect system. For comparison, the experimental band gap of GaAs is in the region of 1.4 eV. The narrowing seen in all cases is to be expected, as DFT is well known to under estimate band gaps. Analysis showed that in the vicinity of the defect, band structure was distorted, with valance states being raised in energy and conduction states lowered. The net result of these two effects was the almost complete closure of the (PBE) band gap. While small band gaps are easier to close, we believe that the effect seen here is robust, as the changes in the LDoS arising from the step defect are very profound. Figure 7.4 shows the change in DoS between a perfect cell and one containing a pair of step defects, the closure of the gap being the most clearly visible result.

7.3.1 Projected DoS

Projected densities of state were computed using the OPTADOS tool. The changes in the band edges was seen to be due largely to the core regions of the

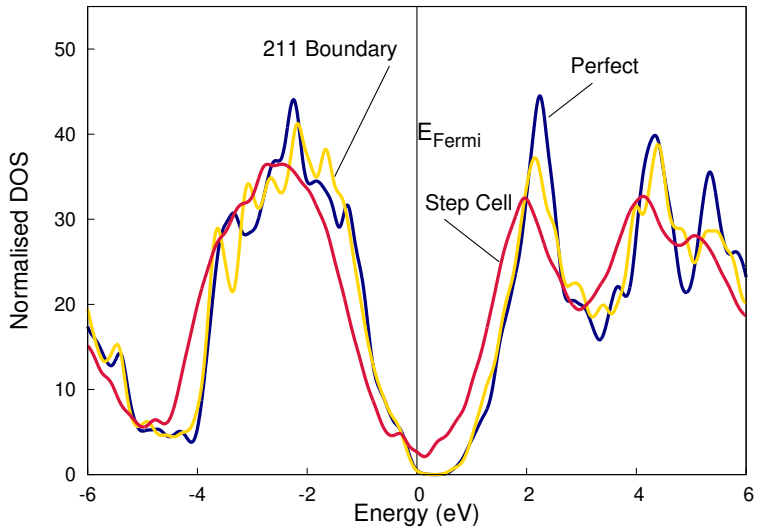


Figure 7.4: The densities of state for bulk GaAs, a 211 grain boundary and the step defect cell. Only the last of these shows a great change in band structure, but in this case the gap is closed completely.

defect. Both the As core and Ga core show atoms with coordination numbers differing from the bulk value (four). Previous workers have shown that this change in coordination tends to lead to gap narrowing in III-V materials.

7.4 Discussion of Optical Properties

As mentioned in the introduction, bulk GaAs is a direct gap semiconductor, with a band gap energy in the infra-red 1.4 eV. The presence of a simple twin grain boundary was shown to have only a relatively minor impact on the gap and DoS, as shown in Figure 7.5. This indicates that simple twins may not have a large impact on the optical properties of GaAs nanowires. The $\Sigma 3$ step twin, however, is a far more radical structural distortion: atoms in the core region of the defect lose their tetrahedral coordination. Even more significantly, some atoms see a change in their bonding with the formation

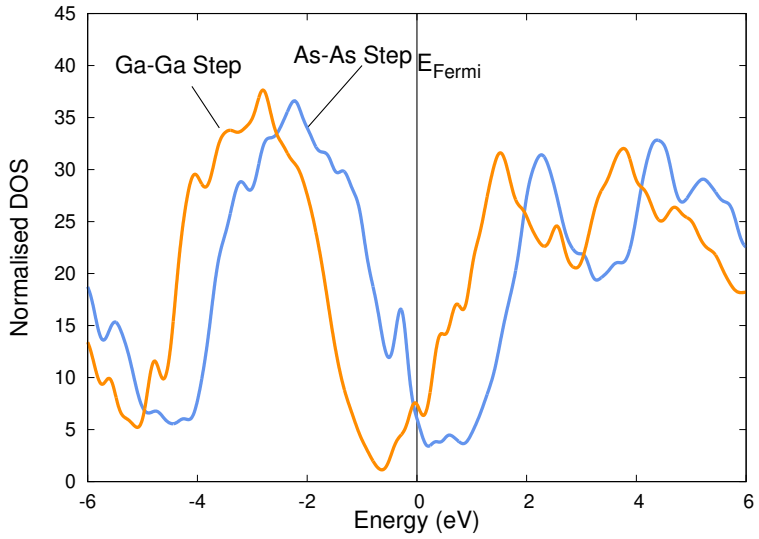


Figure 7.5: The projected densities of state for the As core and Ga core steps. Note that the overlap of these two projected densities is responsible for the gap closure seen in Figure 7.4.

of Ga-Ga and As-As bonds. These changes have a profound impact on the band structure. This narrowing of the bands will lead to a direct effect on both optical absorption and emission frequencies. Furthermore, the states associated with the defect are rather localised. This raises the possibility of them acting as recombination centres.

7.5 Segregation of Phosphorus

The experimentally produced nano wires were doped with 5% phosphorus. A brief study was conducted to assess the segregation of phosphorus to the defect. It was found that phosphorus is favoured in the bulk region with respect to the point of a heptagon in the step defect region by 0.14 eV. The bulk and As-As bond locations were almost identically favoured (difference of 0.02 eV).

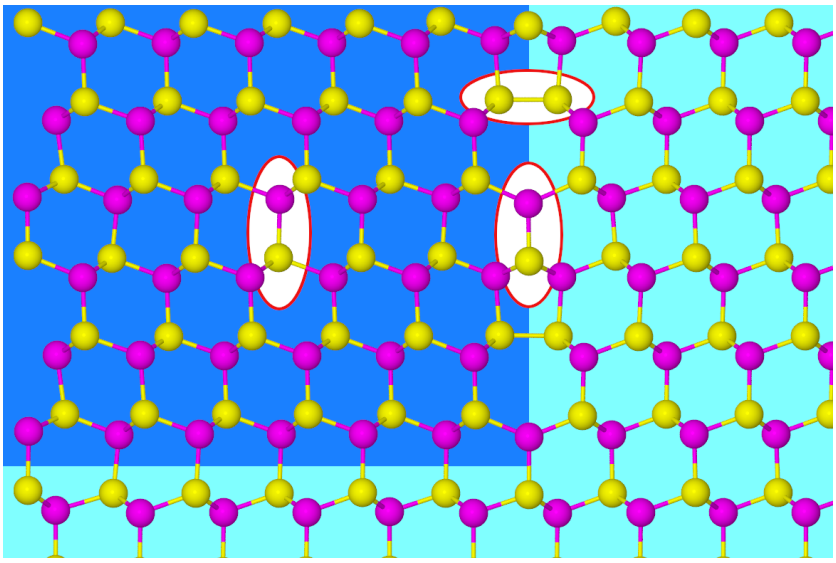


Figure 7.6: The locations of the atoms selected for EELS analysis. On the left is the bulk site. At the top in the middle is the defect core containing the homoatomic bond. Finally in the lower middle is the heptagon-pentagon junction.

7.6 Electron Energy Loss Spectroscopy

This defect is extended in one direction. As a result, it is an ideal candidate for a test of the ideas discussed at the end of Chapter 3: If the TEM beam strikes a column of atoms in the same environment in a material, the EELS signal from those atoms will combine and lead to a clearer signal than would be achieved if only a single defect atom was selected. To assess if the defect would be readily identifiable using EEL spectroscopy, EEL spectra were computed for several Ga and As atoms, both within the defect core and in the bulk-like regions.

Figure 7.7 shows the results of EELS simulations. It is clear that despite quite different coordination environments, there is only a relatively minor difference between the spectra from the defect and bulk-like sites. It is un-

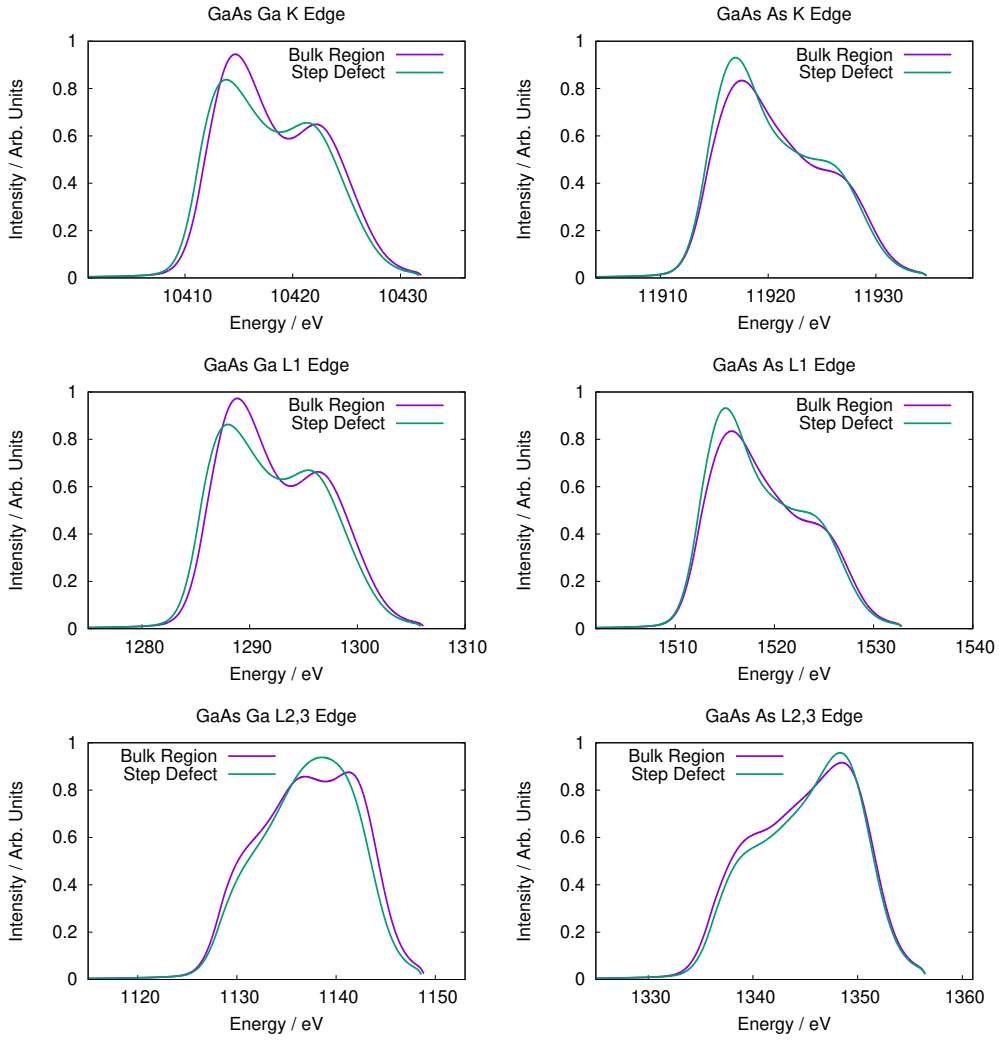


Figure 7.7: EEL spectra for different edges in the GaAs twin defect. The positions of the atoms investigated are indicated in Figure 7.6. The bulk sites are those on the left of that figure while the defect sites are those directly in the middle.

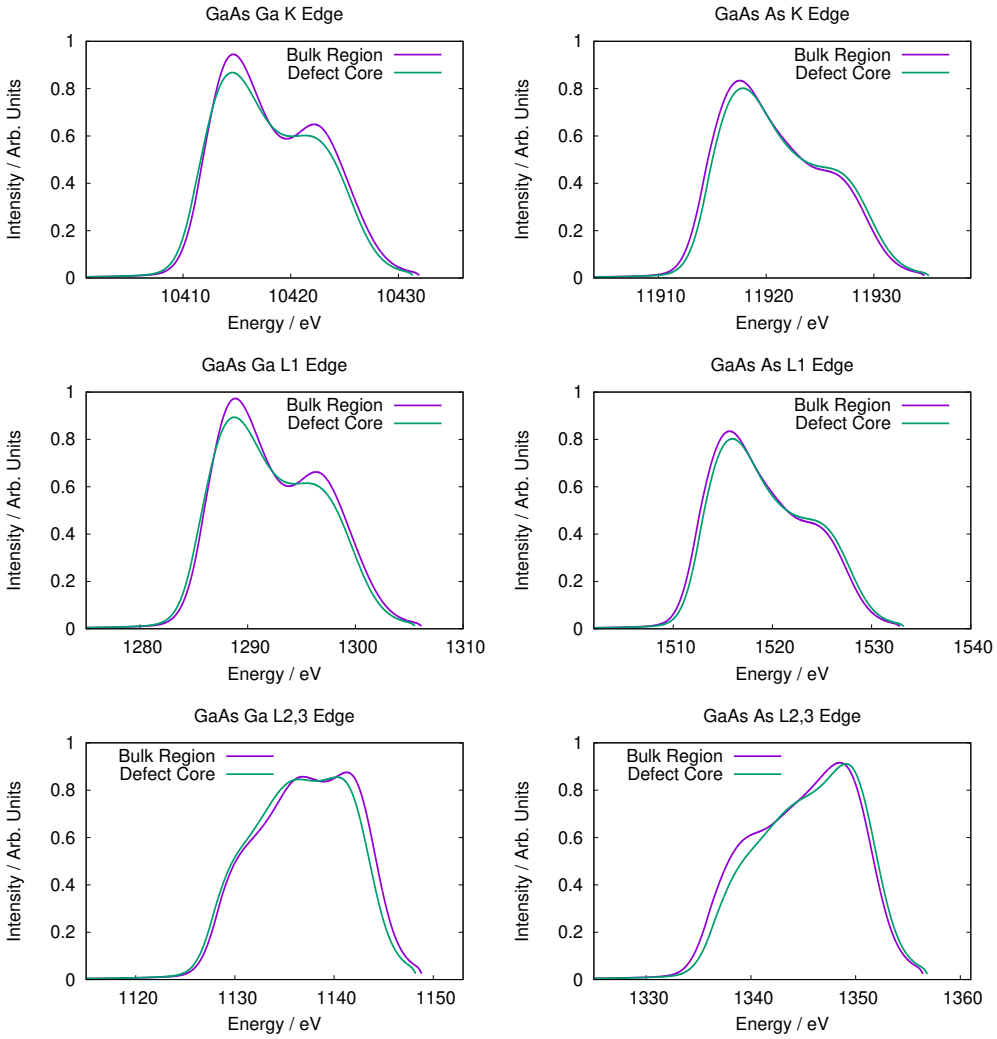


Figure 7.8: EEL spectra for core region (As-As and Ga-Ga bonds). As before, the bulk atomic sites are on the left of figure 7.6. The As-As core site is circled in the top centre of that figure. The Ga-Ga site is equivalent.

likely that for these atoms the defect site will be readily distinguishable from bulk-like GaAs. The only exception to this may be the gallium L2,3 edge, which shows a single peak at 15 eV, as opposed to the double peak seen in the bulk like region.

Spectra for gallium and arsenic atoms involved in Ga-Ga and As-As bonds respectively are shown in figure 7.8. Edges for the atoms in the defect region are overlain on edges for atoms far from the step defect. In the case of all edges, the difference between the bulk and defect region is minor. For the Ga K edge, the double peak structure is less pronounced in the defect region spectrum. This is also seen in the L1 edge. In the case of As, the defective region spectrum shows a reduction in the shoulder structure seen in the bulk L2,3 edge at around 5 eV.

7.7 Conclusion

The band structure of a complex line defect was analysed using DFT. The defect was found to have a profound impact on the properties of the material. Significant narrowing of the band gap was seen. States associated with the defect core likely to assist in recombination were identified. The application of theoretical methods have enabled us to add value to experimental work by providing insight into the properties of the material which were not directly accessible due to the sample size. The behaviour of phosphorus atoms in the vicinity of the defect has also been investigated and characterised. It was found that phosphorus does not tend to segregate to the step defect and in fact shows a preference for the bulk-like portions of the material. Finally, electron energy loss spectra for atoms close to and far from the defect were prepared. Despite the changes in coordination seen in the core region of the defect, there is little change in the predicted EEL spectra. This suggests that EELS may not be the best tool for studying the twin and step defects in GaAs.

The results obtained through simulation have proven useful in rationalising the behaviour of a novel defect, first observed experimentally. Initially, simulation demonstrated the stability of the geometry inferred from TEM data. Analysis of computed band structure has allowed predictions to be made about the impact of this class of defect on the performance of electronic devices made from GaAs. Both of these achievements demonstrate that close integration between theory and experiment can yield important insights into complex nanostructured materials.

Chapter 8

Conclusions

The objective of this work has been to develop and apply methods for understating defects in oxide materials. The basic theory chosen was DFT, owing to its wide adoption and successful application in oxide materials. In Chapter 3, functionality for simulating Electron Energy Loss Spectroscopy was added to the ONETEP code. This functionality was robustly tested against a variety of benchmarks. Initial tests against a second plane wave pseudo potential DFT code demonstrated that the method had been correctly implemented. Testing using an all-electron DFT code for comparison established that the approximations used in implementing the EELS simulation in a PAW code were well justified. Finally, tests with a wide variety of semiconductors and insulators were made, using experimental spectra as a comparison. In all cases, good agreement was obtained, once care had been taken to account for core hole interactions and experimental resolution.

Simulations were undertaken to demonstrate that spectra were well converged with respect to a variety of simulation parameters, an essential aspect of any theoretical investigation.

Once the method had been verified, a set of studies were undertaken in anatase to demonstrate its utility. In the first, the EELS signatures of surface oxygen vacancies were probed. It was shown that defect sites gener-

ated different spectra to non-defective areas. The short range of EELS was demonstrated, with atoms even a short distance from the defect site showing no great change in their spectra compared to the perfect surface.

The effects of a finite beam size were also studied, using a simple linear combination method to generate predicted spectra for surfaces of various thicknesses. This method was also used to investigate the likely thickness of an anatase sheet where defect spectral features may be detectable over the background signal of the perfect material. A method from the literature was used in order to ensure that the energy axes of spectra were correctly aligned. It was found that a defect signal could not be reliably recovered from spectra containing a bulk background. As a conclusion of this study, it was decided that a more useful avenue of study would be heteroatom dopants and interstitials. The former were chosen for subsequent investigations.

A search of the literature indicated a number of possible structures for the inclusion of a nitrogen dopant in anatase. Given the great difference in these structures, it was thought that EELS may provide a means for distinguishing between structures. The objective of this study was to verify the work flow proposed in the introduction: proposed structures are generated, then EEL spectra are simulated. Finally, these simulations are compared to experimental results to determine which structure is the most likely. Sadly, experimental EEL spectra for N dopants in anatase are not available, though it is hoped that this work may trigger interest in recording them.

Knowing which defect is the most stable at a given set of Fermi energy and chemical potentials is vital for making predictions about the behaviour of a material. The oxygen vacancy in anatase has been well studied in the bulk and to a lesser extent in the surface. Chapter 6 sought to unite these two regimes by demonstrating that it is possible to converge formation energies to the bulk value by modeling defects at increasing depth into a slab supercell. This work demonstrated the capacity of ONETEP to simulate large systems, with cells with over 1400 atoms being subject to geometry optimisation. It

is believed that these may be the largest supercells used for a DFT study of oxide materials to date.

This work was hampered in some charge states by the tendency of electrons to delocalise away from the defect. This resulted in behaviour which differed markedly between bulk and surface systems. Hybrid functionals were employed in an attempt to address this difficulty. In the case of the fully ionized 2+ defect, this problem was avoided and the formation energy of the defect in the surface was shown to converge to the formation energy of an isolated defect in the bulk at depths in excess of two layers.

Gallium Arsenide is an important semiconductor. Whilst currently used in the bulk, there are a variety of applications envisaged for nanowires made of this material. Collaboration with an experimental group indicated that a previously unexpected class of defect was present in GaAs nanowires grown by chemical vapour deposition. To understand the role these defects may play in optoelectronic applications, a series of DFT calculations were undertaken. Once the band structure of both the perfect and simply twinned material were ascertained, the complex defect cell was studied. This structure was found to lead to significant reduction of the band gap and would have profound effects in any optoelectronic applications of this material.

The preceding chapters demonstrate that DFT, particularly when used for spectral prediction, is a powerful tool which can be applied to a wide range of semiconductor materials. Furthermore, the spectral prediction capacity added to ONETEP has been shown to produce spectra which closely resemble experimental results. This simplifies the task of comparison and should help increase adoption of the tool by experimentalists wishing to understand their spectra in more detail.

It should be emphasised that two related processes have been demonstrated: First, taking a known defect structure and using it to predict EEL spectral signatures that the defect is expected to produce. The second process takes a series of candidate structures and provides a library of spectra

which may be used to determine which defect exists in a sample.

Work still needs to be done. In the course of this work it became particularly clear that conduction optimisation, whilst indispensable, remains tricky to use. This represents a clear barrier to a greater popularity of the ONETEP code for spectral calculations compared to plane wave codes where low lying conduction states are relatively easy to obtain.

Appendix A

Source Code

A.1 Source Code for the Poisson Solver

```
from __future__ import print_function
import numpy as np
from scipy.fftpack import fftn, fft, \
        ifftn, fftshift, ifftshift
from scipy import linalg
from math import sqrt, exp, pi, erf
from multiprocessing import Pool
from itertools import product

har2ev = 27.211
ang2bohr = 1.889725989

def eps_profile(z,thickness,disp,beta,eps1,eps2):
   ht = thickness / 2.
   return eps1+0.5*(1+erf((z-(ht+disp))/beta)*\
      erf((-(ht+disp)-z)/beta))*(eps2-eps1)
```

```
#@profile
def solve(pts_space=0.8, cell=(20,20,40), verbose=True,
      scale=1., def_pos=None, eps2 = {'x':1.,'y':1.,'z':1.},
      sigma=1., charge=1., thickness=10.,
      eps1 = 1.,beta=2.,disp=2.):

   slab_thickness = thickness*ang2bohr

   x,y,z = map(lambda _x: float(scale)*float(_x),cell)

   x_points = int(2*(int(x/pts_space)//2))
   y_points = int(2*(int(y/pts_space)//2))
   z_points = int(2*(int(z/pts_space)//2))

   x_mp = x_points/2. - 0.5
   x_hp = x_points/2.

   y_mp = y_points/2. - 0.5
   y_hp = y_points/2.

   z_mp = z_points/2. - 0.5
   z_hp = z_points/2.

   norm = charge*(1/(4*pi))*((sqrt(pi)/(4*sigma**1.5))**-1)

   xstep = x/x_points
   ystep = y/y_points
```

```
zstep = z/z_points

Gx = 2*pi/x
Gy = 2*pi/y
Gz = 2*pi/z

#if no defect position is given,
#assume the defect lies in the surface.
if def_pos is None:
    def_pos = slab_thickness / 2.
if def_pos < 0:
    def_pos = slab_thickness / 2. - abs(def_pos)

if verbose:
    print("# Parameters: ")
    print("# points x,y,z", x_points,y_points, z_points)
    print("# x y z", x, y, z)
    print("# xstep ystep zstep", xstep, ystep, zstep)
    print("# sigma", sigma)

rho = np.zeros((x_points,y_points,z_points))
fV = np.zeros((x_points,y_points,z_points),dtype=complex)

#==== Generate the epsilon model ====
if verbose:
    print("#### Generating Epsilon... ####")
model_eps = {}
feps = {}
for axis in 'xyz':
    model_eps[axis] = np.zeros(z_points)
```

```
        for i in range(0,z_points):
            zpos = zstep*(i-z_mp)
            model_eps[axis][i]  = eps_profile(zpos,\
               slab_thickness, disp,beta,eps1, \
               eps2[axis])

        feps[axis] = fft(model_eps[axis])
        feps[axis] = fftshift(feps[axis])/z_points
        feps[axis] = np.append([0.+0.j]*z_points , \
            [feps[axis] , [0.+0.j] * z_points])

    #==== Rho ====
    if verbose:
       print("####␣Generating␣Density...␣####")

    xr = xstep*(np.arange(0,x_points)-x_mp)
    yr = ystep*(np.arange(0,y_points)-y_mp)
    zr = zstep*(np.arange(0,z_points)-z_mp)

    approx_def_idx = int(def_pos/zstep+z_hp)
    dbw_x  = int(sqrt(20/((xstep**2)*sigma)))
    dbw_y  = int(sqrt(20/((ystep**2)*sigma)))
    for k in range(0,z_points):
       for i in range(max(0,x_points//2-dbw_x),\
          min(x_points//2+dbw_x,x_points)):
          for j in range(max(0,y_points//2-dbw_y),\
             min(y_points//2+dbw_y,y_points)):
             rho[i,j,k] =  norm*np.exp(\
                 -sigma*(xr[i]**2+yr[j]**2+(zr[k]-def_pos)**2))
```

```
assert(abs(np.sum(rho)*xstep*ystep*zstep-charge)<0.01)
if verbose:
   print("#### FFT of density ... ####")
frho = fftn(rho)
frho = fftshift(frho)
if verbose:
   print("#### Generating Z matrix ... ####")
eps_mat_x  = np.empty((z_points,z_points),dtype=complex)
eps_mat_y  = np.empty((z_points,z_points),dtype=complex)
eps_mat_z  = np.empty((z_points,z_points),dtype=complex)
Gvec_mat = np.empty((z_points,z_points),dtype=complex)
for k in range(0,z_points):
    for l in range(0,z_points):
        delta = k-l
        eps_idx = delta+z_hp
        epsv_x = feps['x'][int(z_points+eps_idx)]
        epsv_y = feps['y'][int(z_points+eps_idx)]
        epsv_z = feps['z'][int(z_points+eps_idx)]
        Gvec_mat[k][l] = (Gz**2)*(k-z_hp)*(l-z_hp)
        eps_mat_x[k][l]  =  epsv_x
        eps_mat_y[k][l]  =  epsv_y
        eps_mat_z[k][l]  =  epsv_z

if verbose:
   print("#### Computing Fourier Potential ... ####")
mat = np.empty((z_points,z_points),dtype=complex)
#c_inv = None
z_comp_mat = Gvec_mat*eps_mat_z
for i in range(0, x_points):
    for j in range(0, y_points):
```

```
            G_x_sq = (Gx*(i-x_hp))**2
            G_y_sq = (Gy*(j-y_hp))**2
            mat = z_comp_mat + G_x_sq*eps_mat_x +\
                    G_y_sq*eps_mat_y
            if i == x_hp and j == y_hp:
                # to make invertable
                mat[int(z_hp)][int(z_hp)] = 1.
            try:
                fV[i,j,...] = np.linalg.solve(\
                    mat,frho[i,j,...])*4*pi
            except:
                print(fV.shape,frho.shape)
                print(x_points,i,y_points,j,z_points)
                raise
    if verbose:
        print("####␣Inverse␣Fourier␣Transforming...␣####")

    V = ifftn(ifftshift(fV))
    #Average pot. is zero
    V = V - np.sum(V)/(x_points*y_points*z_points)
    #Divide by two to remove double counting
    V *= har2ev
    if verbose:
        print("#␣Method␣2:", e_tot2*har2ev)
    e_tot = np.sum(np.multiply(V,rho))*xstep*ystep*zstep/2.
    if verbose:
        print("####␣Results␣####")
        print("#␣Total␣Energy:",\
            e_tot.real)
        print("#␣Real(total␣Energy)/Imag(Total␣Energy)␣:",\
```

```
            e_tot.real/e_tot.imag)
        print("# GNUPlot-able potential along z axis\
 in middle of simulation cell:")
        print("# Point real(V) imag(V) rho eps")
        Vav = V.sum(axis=0).sum(axis=0)/(x_points*y_points)
        qav = rho.sum(axis=0).sum(axis=0)*xstep*ystep
        for i in range(0,z_points):
            print(i, zstep*(i-z_mp), Vav[i].real, qav[i] ,\
            model_eps['x'][i],V[x_points//2,\
            y_points//2,i].real, y_points//2,i].real)
        print("#",Vav.min().real,Vav.max().real)
    assert(abs(e_tot.real/e_tot.imag)>1e10)
    return e_tot.real
```

Bibliography

(1) Yuangpho, N.; Le, S.; Treerujiraphapong, T.; Khanitchaidecha, W.; Nakaruk, A. *Physica E: Low-dimensional Systems and Nanostructures* **2015**, *67*, 18 –22.

(2) Gautam, A.; Kshirsagar, A.; Biswas, R.; Banerjee, S.; Khanna, P. K. *RSC Advances* **2016**, *6*, 2746–2759.

(3) Ku, Y.; Leu, R.-M.; Lee, K.-C. *Water Research* **1996**, *30*, 2569–2578.

(4) Li, J.; Singh, V. V.; Sattayasamitsathit, S.; Orozco, J.; Kaufmann, K.; Dong, R.; Gao, W.; Jurado-Sanchez, B.; Fedorak, Y.; Wang, J. *ACS Nano* **2014**, *8*, 11118–11125.

(5) Asahi, R.; Morikawa, T.; Ohwaki, T.; Aoki, K.; Taga, Y. *Science* **2001**, *293*, 269–271.

(6) Umebayashi, T; Yamaki, T; Itoh, H; Asai, K *Applied Physics Letters* **2002**, *81*, 454–456.

(7) Irie, H.; Watanabe, Y.; Hashimoto, K. *Chemistry Letters* **2003**, *32*, 772–773.

(8) Williams, D. B.; Carter, C. B. In *Transmission electron microscopy*; Springer: 1996, pp 3–17.

(9) Egerton, R. F., *Electron energy-loss spectroscopy in the electron microscope*; Springer Science & Business Media: 2011.

(10) Tiemeijer, P. In *Inst. Phys. Conf. Ser*, 1999; Vol. 161, pp 191–194.

(11) Liu, G.; Wang, Y.; Zou, B.; Liang, W.; Alford, N. M.; McComb, D. W.; Petrov, P. K. *The Journal of Physical Chemistry C* **2016**, *120*, 16681–16686.

(12) Nicholls, R. J.; Murdock, A. T.; Tsang, J.; Britton, J.; Pennycook, T. J.; Koos, A.; Nellist, P. D.; Grobert, N.; Yates, J. R. *ACS nano* **2013**, *7*, 7145–7150.

(13) Ramasse, Q. M.; Seabourne, C. R.; Kepaptsoglou, D.-M.; Zan, R.; Bangert, U.; Scott, A. J. *Nano letters* **2013**, *13*, 4989–4995.

(14) Chang, K.; Nuhfer, N.; Porter, L.; Wahab, Q *Applied Physics Letters* **2000**, *77*, 2186–2188.

(15) Scheu, C; Dehm, G; Rühle, M; Brydson, R *Philosophical Magazine A* **1998**, *78*, 439–465.

(16) Brydson, R; Sauer, H; Engel, W; Thomass, J.; Zeitler, E; Kosugi, N; Kuroda, H *Journal of Physics: Condensed Matter* **1989**, *1*, 797.

(17) Laffont, L; Delacourt, C; Gibot, P; Wu, M. Y.; Kooyman, P; Masquelier, C; Tarascon, J. M. *Chemistry of Materials* **2006**, *18*, 5520–5529.

(18) Griffiths, I.; Cherns, D; Albert, S; Bengoechea-Encabo, A; Angel Sanchez, M; Calleja, E; Schimpke, T; Strassburg, M *Journal of microscopy* **2016**, *262*, 167–170.

(19) Liebscher, C.; Freysoldt, C; Dennenwaldt, T; Harzer, T.; Dehm, G *Ultramicroscopy* **2017**, *178*, 96–104.

(20) Goode, A.; Hine, N.; Chen, S; Bergin, S.; Shaffer, M.; Ryan, M.; Haynes, P.; Porter, A.; McComb, D. *Chemical Communications* **2014**, *50*, 6744–6747.

(21) Rehr, J. J.; Kas, J. J.; Prange, M. P.; Sorini, A. P.; Takimoto, Y.; Vila, F. *Comptes Rendus Physique* **2009**, *10*, 548–559.

(22) Rehr, J. J.; Kas, J. J.; Vila, F. D.; Prange, M. P.; Jorissen, K. *Physical Chemistry Chemical Physics* **2010**, *12*, 5503–5513.

(23) Bethe, H. *Annalen der Physik* **1930**, *397*, 325–400.

(24) Leapman, R.; Fejes, P.; Silcox, J *Physical review B* **1983**, *28*, 2361.

(25) Rez, P.; Bruley, J.; Brohan, P.; Payne, M.; Garvie, L. *Ultramicroscopy* **1995**, *59*, Proceedings of the 2nd international workshop on Electron Energy Loss Spectroscopy and Imaging, 159 –167.

(26) Olovsson, W; Tanaka, I; Mizoguchi, T; Puschnig, P; Ambrosch-Draxl, C *Physical Review B* **2009**, *79*, 041102.

(27) Ikeno, H.; Mizoguchi, T.; Koyama, Y.; Kumagai, Y.; Tanaka, I. *Ultramicroscopy* **2006**, *106*, Proceedings of the International Workshop on Enhanced Data Generated by Electrons, 970 –975.

(28) Mizoguchi, T.; Olovsson, W.; Ikeno, H.; Tanaka, I. *Micron* **2010**, *41*, 695 –709.

(29) Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864.

(30) Oliveira, M. Functionals list unreleased. `https://gitlab.com/libxc/libxc/wikis/Functionals-list-unreleased` (accessed 07/24/2017).

(31) Perdew, J. P.; Schmidt, K. **2001**, *577*, 1–20.

(32) Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133–A1138.

(33) Dirac, P. A. M. *Mathematical Proceedings of the Cambridge Philosophical Society* **1930**, *26*, 376–385.

(34) Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048–5079.

(35) Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* **1980**, *45*, 566–569.

(36) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. *Rev. Mod. Phys.* **1992**, *64*, 1045–1097.

(37) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical review letters* **1996**, *77*, 3865.

(38) Perdew, J. P.; Ruzsinszky, A.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E.; Constantin, L. A.; Zhou, X.; Burke, K. *Phys. Rev. Lett.* **2008**, *100*, 136406.

(39) Zhang, Y.; Yang, W. *Phys. Rev. Lett.* **1998**, *80*, 890–890.

(40) Klimeš, J.; Bowler, D. R.; Michaelides, A. *Journal of Physics: Condensed Matter* **2010**, *22*, 022201.

(41) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Physical Review Letters* **2003**, *91*, 146401.

(42) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. *The Journal of Physical Chemistry* **1994**, *98*, 11623–11627.

(43) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *The Journal of Chemical Physics* **2003**, *118*, 8207–8215.

(44) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. *The Journal of Chemical Physics* **2006**, *124*, 219906.

(45) Krukau, A. V.; Vydrov, O. A.; Izmaylov, A. F.; Scuseria, G. E. *The Journal of Chemical Physics* **2006**, *125*, 224106.

(46) Cohen, M. H.; Heine, V. *Phys. Rev.* **1961**, *122*, 1821–1826.

(47) Hamann, D. R.; Schlüter, M.; Chiang, C. *Phys. Rev. Lett.* **1979**, *43*, 1494–1497.

(48) Bachelet, G. B.; Schlüter, M. *Phys. Rev. B* **1982**, *25*, 2103–2108.

(49) Denteneer, P. J. H.; van Haeringen, W *Journal of Physics C: Solid State Physics* **1985**, *18*, 4127.

(50) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892–7895.

(51) Blöchl, P. E. *Physical Review B* **1994**, *50*, 17953–17979.

(52) Castro, A.; Appel, H.; Oliveira, M.; Rozzi, C. A.; Andrade, X.; Lorenzen, F.; Marques, M. A.; Gross, E.; Rubio, A. *physica status solidi (b)* **2006**, *243*, 2465–2488.

(53) Goedecker, S. *Rev. Mod. Phys.* **1999**, *71*, 1085–1123.

(54) Kohn, W. *Phys. Rev. Lett.* **1996**, *76*, 3168–3171.

(55) Prodan, E.; Kohn, W. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 11635–11638.

(56) Brouder, C.; Panati, G.; Calandra, M.; Mourougane, C.; Marzari, N. *Phys. Rev. Lett.* **2007**, *98*, 046402.

(57) Mohr, S.; Ratcliff, L. E.; Boulanger, P.; Genovese, L.; Caliste, D.; Deutsch, T.; Goedecker, S. *The Journal of Chemical Physics* **2014**, *140*, 204110.

(58) Gillan, M.; Bowler, D.; Torralba, A.; Miyazaki, T. *Computer Physics Communications* **2007**, *177*, Proceedings of the Conference on Computational Physics 2006, 14 –18.

(59) Liang, W.; Saravanan, C.; Shao, Y.; Baer, R.; Bell, A. T.; Head-Gordon, M. *The Journal of chemical physics* **2003**, *119*, 4117–4125.

(60) Lin, L.; García, A.; Huhs, G.; Yang, C. *Journal of Physics: Condensed Matter* **2014**, *26*, 305503.

(61) Seewald, P. Enabling DFT Simulations of Large Metallic Systems by Integrating the PEXSI Method into CP2K., Ph.D. Thesis, 2015.

(62) Yang, W.; Lee, T.-S. *The Journal of chemical physics* **1995**, *103*, 5674–5678.

(63) Nakai, H.; Kobayashi, M. *Procedia Computer Science* **2011**, *4*, 1145–1150.

(64) Bowler, D. R.; Miyazaki, T *Reports on Progress in Physics* **2012**, *75*, 036503.

(65) Hernández, E.; Gillan, M. J.; Goringe, C. M. *Physical Review B* **1996**, *53*, 7147–7157.

(66) Skylaris, C.-K.; Haynes, P. D.; Mostofi, A. A.; Payne, M. C. *The Journal of Chemical Physics* **2005**, *122*.

(67) Mostofi, A. A.; Skylaris, C.-K.; Haynes, P. D.; Payne, M. C. *Computer Physics Communications* **2002**, *147*, 788 –802.

(68) Hotelling, H. *The Annals of Mathematical Statistics* **1943**, *14*, 1–34.

(69) Haynes, P. D.; Skylaris, C.-K.; Mostofi, A. A.; Payne, M. C. *Journal of Physics: Condensed Matter* **2008**, *20*, 294207.

(70) Palser, A. H. R.; Manolopoulos, D. E. *Phys. Rev. B* **1998**, *58*, 12704–12711.

(71) Haynes, P. D.; Payne, M. C. *Phys. Rev. B* **1999**, *59*, 12173–12176.

(72) Li, X.-P.; Nunes, R. W.; Vanderbilt, D. *Phys. Rev. B* **1993**, *47*, 10891–10894.

(73) Nunes, R. W.; Vanderbilt, D. *Phys. Rev. B* **1994**, *50*, 17611–17614.

(74) McWeeny, R. *Rev. Mod. Phys.* **1960**, *32*, 335–369.

(75) Haynes, P.; Payne, M. *Solid State Communications* **1998**, *108*, 737 –741.

(76) Hine, N. D.; Haynes, P. D.; Mostofi, A. A.; Skylaris, C.-K.; Payne, M. C. *Computer Physics Communications* **2009**, *180*, 1041–1053.

(77) Wilkinson, K. A.; Hine, N. D.; Skylaris, C.-K. *Journal of chemical theory and computation* **2014**, *10*, 4782–4794.

(78) Skylaris, C.-K.; Mostofi, A. A.; Haynes, P. D.; Diéguez, O.; Payne, M. C. *Phys. Rev. B* **2002**, *66*, 035119.

(79) Haynes, P. D.; Mostof, A. A.; Skylaris, C.-K.; Payne, M. C. *Journal of Physics: Conference Series* **2006**, *26*, 143.

(80) Ratcliff, L. E.; Hine, N. D. M.; Haynes, P. D. *Physical Review B* **2011**, *84*, 165131–165141.

(81) Hine, N. D. M. *Journal of Physics: Condensed Matter* **2017**, *29*, 024001.

(82) Zuehlsdorff, T. J.; Haynes, P. D.; Payne, M. C.; Hine, N. D. M. *The Journal of Chemical Physics* **2017**, *146*, 124504.

(83) Verga, L. G.; Aarons, J.; Sarwar, M.; Thompsett, D.; Russell, A. E.; Skylaris, C.-K. *Phys. Chem. Chem. Phys.* **2016**, *18*, 32713–32722.

(84) Cole, D. J.; Hine, N. D. M. *Journal of Physics: Condensed Matter* **2016**, *28*, 393001.

(85) Bell, R. A.; Payne, M. C.; Mostofi, A. A. *The Journal of Chemical Physics* **2014**, *141*, 164703.

(86) Constantinescu, G. C.; Hine, N. D. M. *Nano Letters* **2016**, *16*, PMID: 27028122, 2586–2594.

(87) Tait, E. W.; Ratcliff, L. E.; Payne, M. C.; Haynes, P. D.; Hine, N. D. M. *Journal of Physics: Condensed Matter* **2016**, *28*, 195202.

(88) Lindner, T.; Sauer, H.; Engel, W.; Kambe, K. *Phys. Rev. B* **1986**, *33*, 22–24.

(89) Korringa, J *Physica* **1947**, *13*, 392 –400.

(90) Kohn, W.; Rostoker, N. *Phys. Rev.* **1954**, *94*, 1111–1120.

(91) Taillefumier, M.; Cabaret, D.; Flank, A.-M.; Mauri, F. *Phys. Rev. B* **2002**, *66*, 195107.

(92) Gao, S.-P.; Pickard, C. J.; Perlov, A.; Milman, V. *Journal of Physics: Condensed Matter* **2009**, *21*, 104203.

(93) Aguiar, J. A.; Grönbech-Jensen, N; Perlov, A; Milman, V; Gao, S. P.; Pickard, C. J.; Browning, N. D. *Journal of Physics: Conference Series* **2010**, *241*, 012062.

(94) Seabourne, C. R.; Scott, A. J.; Vaughan, G.; Brydson, R.; Wang, S.-G.; Ward, R. C.; Wang, C.; Kohn, A.; Mendis, B.; Petford-Long, A. K. *Ultramicroscopy* **2010**, *110*, Proceedings of the International Workshop On Enhanced Data Generated By Electrons, 1059 –1069.

(95) Guo, N.; Fingland, B. R.; Williams, W. D.; Kispersky, V. F.; Jelic, J.; Delgass, W. N.; Ribeiro, F. H.; Meyer, R. J.; Miller, J. T. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5678–5693.

(96) Bradley, M. K.; Robinson, J.; Woodruff, D. P. *The Journal of Physical Chemistry C* **2013**, *117*, 12591–12599.

(97) Mauri, F.; Car, R. *Physical review letters* **1995**, *75*, 3166.

(98) Gao, S.-P.; Pickard, C. J.; Payne, M. C.; Zhu, J.; Yuan, J. *Physical Review B* **2008**, *77*, 115122.

(99) Lu, J.; Gao, S.-P. *Computational Materials Science* **2013**, *68*, 335 – 341.

(100) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. J.; Refson, K.; Payne, M. *Z. Kristall.* **2005**, *220*, 567–570.

(101) Duscher, G.; Buczko, R.; Pennycook, S.; Pantelides, S. *Ultramicroscopy* **2001**, *86*, International Symposium on Spectroscopy of Materials, 355 –362.

(102) Donval, G.; Moreau, P.; Danet, J.; Larbi, S. J.-S.; Bayle-Guillemaud, P.; Boucher, F. *Physical Chemistry Chemical Physics* **2017**, *19*, 1320–1327.

(103) Morris, A. J.; Nicholls, R. J.; Pickard, C. J.; Yates, J. R. *Computer Physics Communications* **2014**, *185*, 1477–1485.

(104) Nicholls, R.; Morris, A.; Pickard, C.; Yates, J. In *Journal of Physics: Conference Series*, 2012; Vol. 371, p 012062.

(105) Elsässer, C.; Köstlmeier, S. *Ultramicroscopy* **2001**, *86*, International Symposium on Spectroscopy of Materials, 325 –337.

(106) Mizoguchi, T.; Tanaka, I.; Gao, S.-P.; Pickard, C. J. *Journal of Physics: Condensed Matter* **2009**, *21*, 104204.

(107) Fletcher, D. A.; McMeeking, R. F.; Parkin, D. *Journal of Chemical Information and Computer Sciences* **1996**, *36*, 746–749.

(108) Jollet, F.; Torrent, M.; Holzwarth, N. *Computer Physics Communications* **2014**, *185*, 1246 –1254.

(109) Lejaeghere, K.; Speybroeck, V. V.; Oost, G. V.; Cottenier, S. *Critical Reviews in Solid State and Materials Sciences* **2014**, *39*, 1–24.

(110) Dewhurst, K.; Sharma, S.; Nordström, L.; Cricchio, F.; Gråsäs, O.; Gross, H. Elk Code. `http://elk.sourceforge.net/` (accessed 08/01/2017).

(111) Lian, C.-S.; Wang, X.-Q.; Wang, J.-T. *The Journal of Chemical Physics* **2013**, *138* 024702, –.

(112) Trucano, P.; Chen, R. *Nature* **1975**.

(113) Bernuy-Lopez, C.; Allix, M.; Bridges, C. A.; Claridge, J. B.; Rosseinsky, M. J. *Chemistry of materials* **2007**, *19*, 1035–1043.

(114) Többens, D.; Stüßer, N; Knorr, K; Mayer, H.; Lampert, G In *Materials Science Forum*, 2001; Vol. 378, pp 288–293.

(115) Gao, S.-P. *Solid State Communications* **2012**, *152*, 1817 –1820.

(116) Wen, B.; Zhao, J.; Melnik, R.; Tian, Y. *Phys. Chem. Chem. Phys.* **2011**, *13*, 14565–14570.

(117) Ewels, P.; Sikora, T.; Serin, V.; Ewels, C. P.; Lajaunie, L. *Microscopy and Microanalysis* **2016**, *FirstView*, 1–8.

(118) Hamon, A.-L.; Verbeeck, J.; Schryvers, D.; Benedikt, J.; M. C. M. v. d. Sanden, R. *J. Mater. Chem.* **2004**, *14*, 2030–2035.

(119) Cheng, H.; Selloni, A. *Phys. Rev. B* **2009**, *79*, 092101.

(120) Li, H.; Guo, Y.; Robertson, J. *The Journal of Physical Chemistry C* **2015**, *119*, 18160–18166.

(121) Zhu, G.-Z.; Botton, G. A. *Microscopy and Microanalysis* **2014**, *20*, 649–657.

(122) Brydson, R; Sauer, H; Engel, W; Thomass, J.; Zeitler, E; Kosugi, N; Kuroda, H *Journal of Physics: Condensed Matter* **1989**, *1*, 797.

(123) Stoyanov, E; Langenhorst, F; Steinle-Neumann, G *American Mineralogist* **2007**, *92*, 577–586.

(124) Bosman, M; Watanabe, M; Alexander, D.; Keast, V. *Ultramicroscopy* **2006**, *106*, 1024–1032.

(125) Zhang, Y.; Guerra-Nuñez, C.; Li, M.; Michler, J.; Park, H. G.; Rossell, M. D.; Erni, R.; Utke, I. *Chemistry of Materials* **2016**, *28*, 3488–3496.

(126) Cocchi, C.; Zschiesche, H.; Nabok, D.; Mogilatenko, A.; Albrecht, M.; Galazka, Z.; Kirmse, H.; Draxl, C.; Koch, C. T. *Phys. Rev. B* **2016**, *94*, 075147.

(127) Laffont, L; Delacourt, C; Gibot, P; Wu, M. Y.; Kooyman, P; Masquelier, C; Tarascon, J. M. *Chemistry of Materials* **2006**, *18*, 5520–5529.

(128) Bugnet, M.; Radtke, G.; Woo, S. Y.; Zhu, G.-z.; Botton, G. A. *Phys. Rev. B* **2016**, *93*, 020102.

(129) Gázquez, J.; Sánchez-Santolino, G.; Biškup, N.; Roldán, M. A.; Cabero, M; Pennycook, S. J.; Varela, M. *Materials Science in Semiconductor Processing* **2017**, *65*, 49–63.

(130) Ramasse, Q. M.; Seabourne, C. R.; Kepaptsoglou, D.-M.; Zan, R.; Bangert, U.; Scott, A. J. *Nano letters* **2013**, *13*, 4989–4995.

(131) Hine, N. D. M.; Robinson, M.; Haynes, P. D.; Skylaris, C.-K.; Payne, M. C.; Mostofi, A. A. *Phys. Rev. B* **2011**, *83*, 195102.

(132) Ewels, P.; Sikora, T.; Serin, V.; Ewels, C. P.; Lajaunie, L. *Microscopy and Microanalysis* **2016**, *FirstView*, 1–8.

(133) Mosquera, A. A.; Endrino, J. L.; Albella, J. M. *J. Anal. At. Spectrom.* **2014**, *29*, 736–742.

(134) Thomas, A.; Flavell, W.; Mallick, A.; Kumarasinghe, A.; Tsoutsou, D; Khan, N; Chatwin, C; Rayner, S; Smith, G.; Stockbauer, R., et al. *Physical Review B* **2007**, *75*, 035105.

(135) De la Peña, F.; Berger, M.-H.; Hochepied, J.-F.; Dynys, F.; Stephan, O.; Walls, M. *Ultramicroscopy* **2011**, *111*, 169 –176.

(136) Cong, Y.; Zhang, J.; Chen, F.; Anpo, M. *The Journal of Physical Chemistry C* **2007**, *111*, 6976–6982.

(137) Batzill, M.; Morales, E. H.; Diebold, U. *Physical review letters* **2006**, *96*, 026103.

(138) Chen, Q.; Tang, C.; Zheng, G. *Physica B: Condensed Matter* **2009**, *404*, 1074–1078.

(139) Ihara, T.; Miyoshi, M.; Iriyama, Y.; Matsumoto, O; Sugihara, S *Applied Catalysis B: Environmental* **2003**, *42*, 403–409.

(140) Chen, H.; Dawson, J. A. *The Journal of Physical Chemistry C* **2015**, *119*, 15890–15895.

(141) Zhao, Z.; Liu, Q. *Journal of Physics D: Applied Physics* **2007**, *41*, 025105.

(142) Yang, K.; Dai, Y.; Huang, B. *The Journal of Physical Chemistry C* **2007**, *111*, 12086–12090.

(143) Reyes-Garcia, E. A.; Sun, Y.; Reyes-Gil, K.; Raftery, D. *The Journal of Physical Chemistry C* **2007**, *111*, 2738–2748.

(144) Vittadini, A; Selloni, A; Rotzinger, F.; Grätzel, M *Physical Review Letters* **1998**, *81*, 2954.

(145) Vittadini, A; Selloni, A; Rotzinger, F.; Grätzel, M *The Journal of Physical Chemistry B* **2000**, *104*, 1300–1306.

(146) Chen, Q. L.; Li, B.; Zheng, G.; He, K. H.; Zheng, A. S. *Physica B: Condensed Matter* **2011**, *406*, 3841–3846.

(147) Zhang, W.; Yin, J.-R.; Tang, X.-Q.; Zhang, P.; Ding, Y.-H. *Physica E: Low-dimensional Systems and Nanostructures* **2017**, *85*, 259–263.

(148) Setvín, M.; Aschauer, U.; Scheiber, P.; Li, Y.-F.; Hou, W.; Schmid, M.; Selloni, A.; Diebold, U. *Science* **2013**, *341*, 988–991.

(149) Scheiber, P.; Fidler, M.; Dulub, O.; Schmid, M.; Diebold, U.; Hou, W.; Aschauer, U.; Selloni, A. *Phys. Rev. Lett.* **2012**, *109*, 136103.

(150) He, Y.; Dulub, O.; Cheng, H.; Selloni, A.; Diebold, U. *Phys. Rev. Lett.* **2009**, *102*, 106105.

(151) Na-Phattalung, S.; Smith, M. F.; Kim, K.; Du, M.-H.; Wei, S.-H.; Zhang, S.; Limpijumnong, S. *Physical Review B* **2006**, *73*, 125205.

(152) Morgan, B. J.; Watson, G. W. *The Journal of Physical Chemistry C* **2010**, *114*, 2321–2328.

(153) Lazzeri, M.; Vittadini, A.; Selloni, A. *Phys. Rev. B* **2001**, *63*, 155409.

(154) Haa, M.-A.; Alexandrova, A. N. *Journal of Chemical Theory and Computation* **2016**, *12*, PMID: 27163165, 2889–2895.

(155) Stetsovych, O.; Todorović, M.; Shimizu, T. K.; Moreno, C.; Ryan, J. W.; León, C. P.; Sagisaka, K.; Palomares, E.; Matolín, V.; Fujita, D., et al. *Nature communications* **2015**, *6*.

(156) Li, Y.; Gao, Y. *Phys. Rev. Lett.* **2014**, *112*, 206101.

(157) Cheng, H.; Selloni, A. *The Journal of chemical physics* **2009**, *131*, 054703.

(158) Yoon, S. D.; Chen, Y.; Yang, A.; Goodrich, T. L.; Zuo, X.; Arena, D. A.; Ziemer, K.; Vittoria, C.; Harris, V. G. *Journal of Physics: Condensed Matter* **2006**, *18*, L355.

(159) Zhou, S.; Potzger, K.; Krause, M.; Talut, G.; Helm, M.; Fassbender, J.; Zvyagin, S. A.; Wosnitza, J.; Schmidt, H. *Phys. Rev. B* **2009**, *79*, 113201.

(160) Yang, K.; Dai, Y.; Huang, B.; Feng, Y. P. *Phys. Rev. B* **2010**, *81*, 033202.

(161) Finazzi, E.; Di Valentin, C.; Pacchioni, G.; Selloni, A. *The Journal of chemical physics* **2008**, *129*, 154113.

(162) Janotti, A; Varley, J.; Rinke, P; Umezawa, N; Kresse, G; Van de Walle, C. *Physical Review B* **2010**, *81*, 085212.

(163) Mattioli, G.; Alippi, P.; Filippone, F.; Caminiti, R.; Amore Bonapasta, A. *The Journal of Physical Chemistry C* **2010**, *114*, 21694–21704.

(164) Zhou, S.; Potzger, K.; Krause, M.; Talut, G.; Helm, M.; Fassbender, J.; Zvyagin, S. A.; Wosnitza, J.; Schmidt, H. Rappe Group., Accessed September 2017, 2009.

(165) Turban, D. H. P.; Teobaldi, G.; O'Regan, D. D.; Hine, N. D. M. *Phys. Rev. B* **2016**, *93*, Accessed September 2017, 165102.

(166) Makov, G; Payne, M. *Physical Review B* **1995**, *51*, Accessed September 2017, 4014.

(167) Freysoldt, C.; Neugebauer, J.; Van de Walle, C. G. *Phys. Rev. Lett.* **2009**, *102*, Accessed September 2017, 016402.

(168) Komsa, H.-P.; Pasquarello, A. *Physical review letters* **2013**, *110*, Accessed September 2017, 095505.

(169) Castleton, C. W. M.; Höglund, A.; Mirbt, S. *Phys. Rev. B* **2006**, *73*, Accessed September 2017, 035215.

(170) Hine, N.; Frensch, K; Foulkes, W.; Finnis, M. *Physical Review B* **2009**, *79*, Accessed September 2017, 024112.

(171) Zhang, S. B.; Northrup, J. E. *Phys. Rev. Lett.* **1991**, *67*, Accessed September 2017, 2339–2342.

(172) Finnis, M.; Lozovoi, A.; Alavi, A *Annu. Rev. Mater. Res.* **2005**, *35*, Accessed September 2017, 167–207.

(173) Selçuk, S.; Selloni, A. *The Journal of chemical physics* **2014**, *141*, Accessed September 2017, 084705.

(174) Gonzalez, R. J.; Zallen, R.; Berger, H. *Phys. Rev. B* **1997**, *55*, Accessed September 2017, 7014–7017.

(175) Mikami, M.; Nakamura, S.; Kitao, O.; Arakawa, H. *Phys. Rev. B* **2002**, *66*, Accessed September 2017, 155213.

(176) Giarola, M.; Sanson, A.; Monti, F.; Mariotto, G.; Bettinelli, M.; Speghini, A.; Salviulo, G. *Phys. Rev. B* **2010**, *81*, Accessed September 2017, 174305.

(177) Carlson, B. C. *Journal of Mathematical Physics* **1961**, *2*, Accessed September 2017, 441–450.

(178) Naik, M.; Jain, M. *Phys. Rev. B* **2017**, *79*, Accessed September 2017, 113201.

(179) Lambrecht, W. R. *physica status solidi (b)* **2011**, *248*, Accessed September 2017, 1547–1558.

(180) Boonchun, A.; Reunchan, P.; Umezawa, N. *Physical Chemistry Chemical Physics* **2016**, *18*, Accessed September 2017, 30040–30046.

(181) Deskins, N. A.; Dupuis, M. *Phys. Rev. B* **2007**, *75*, Accessed September 2017, 195212.

(182) Di Valentin, C.; Selloni, A. *The Journal of Physical Chemistry Letters* **2011**, *2*, Accessed September 2017, 2223–2228.

(183) Deák, P.; Kullgren, J.; Frauenheim, T. *physica status solidi (RRL)-Rapid Research Letters* **2014**, *8*, Accessed September 2017, 583–586.

(184) Sanchez, A. M.; Zhang, Y.; Tait, E. W.; Hine, N. D.; Liu, H.; Beanland, R. *Nano Letters* **2017**, *17*, Accessed September 2017, 2454–2459.

(185) Jacobsson, D.; Panciera, F.; Tersoff, J.; Reuter, M. C.; Lehmann, S.; Hofmann, S.; Dick, K. A.; Ross, F. M. *Nature* **2016**, *531*, Accessed September 2017, 317–322.

(186) Colombo, C; Heiß, M; Grätzel, M; Fontcuberta i Morral, A *Applied Physics Letters* **2009**, *94*, Accessed September 2017, 173108.

(187) Åberg, I.; Vescovi, G.; Asoli, D.; Naseem, U.; Gilboy, J. P.; Sundvall, C.; Dahlgren, A.; Svensson, K. E.; Anttu, N.; Björk, M. T., et al. *IEEE Journal of Photovoltaics* **2016**, *6*, Accessed September 2017, 185–190.

(188) Svensson, C. P. T.; Mårtensson, T.; Trägårdh, J.; Larsson, C.; Rask, M.; Hessman, D.; Samuelson, L.; Ohlsson, J. *Nanotechnology* **2008**, *19*, Accessed September 2017, 305201.

(189) Peng, K.; Parkinson, P.; Fu, L.; Gao, Q.; Jiang, N.; Guo, Y.-N.; Wang, F.; Joyce, H. J.; Boland, J. L.; Tan, H. H., et al. *Nano letters* **2014**, *15*, Accessed September 2017, 206–210.

(190) Trukhin, V.; Buyskikh, A.; Kaliteevskaya, N.; Bourauleuv, A.; Samoilov, L.; Samsonenko, Y. B.; Cirlin, G.; Kaliteevski, M.; Gallant, A. *Applied Physics Letters* **2013**, *103*, Accessed September 2017, 072108.